INVESTIGACION y CIENCIA

¿QUE HAY DE CIERTO EN EL CAMBIO DE CLIMA?

EL LEP, ESPERANZA DE LA FISICA EUROPEA

BASE GENETICA DEL DESARROLLO CORPORAL

Edición española de SCIENTIFIC AMERICAN



PIROTECNIA

SEPTIEMBRE 1990
500 PTAS.



Septiembre de 1990 Número 168

SECCIONES

COLABORADORES DE ESTE NUMERO

Asesoramiento y traducción:

Manuel Puigcerver: *El gran debate sobre el clima*; Mª Verónica Etchart: *Genes con homeobox y el plan corporal de los vertebrados*; Ramón Pascual: *El colisionador LEP*; Mª Pilar Ramos: *Origen de la diabetes*; Amando García: *Pirotecnia*; Carmen Gómez-Limón: *Chancro del castaño*; Luis Bou: *Teoría de Ramsey*, *Juegos matemáticos* y *Juegos de ordenador*.

Ciencia y sociedad:

Ramón Pascual, Josep-Enric Llebot y Joandoménec Ros

Ciencia y empresa:

Manuel Puigcerver

Libros:

Ramón Margalef, Manuel García Doncel, Emilio Elizalde y Luis Alonso

PROCEDENCIA DE LAS ILUSTRACIONES

Página	Fuente
5	AP/Wide World Photos
6	Bernard G. Mendosa, NOAA
7-11	Johnny Johnson
15-17	Patricia J. Wynne
18	Edward Bell
19	Patricia J. Wynne
20	G. Oliver (*izq.*), A. Molven, Univ. Oregón (*drcha.*)
23	George Retseck
24	George Retseck (*arriba*), Johnny Johnson (*abajo*)
25-27	George Retseck
28-29	CERN
38	Gabor Kiss
39	William Riley, U. Florida
40	Gabor Kiss
41	Mark A. Atkinson
42	Gabor Kiss
43	G. Kiss (*izq. y drcha.*), D. C. Wiley, Univ. Harvard (*centro*)
44	Dana Burns-Pizer
49-50	NASA/JPL
51	S. S. Limaye
52	A. S. Lavega y Magda Màría
53	T. E. Dowling, A. P. Ingersoll y Magda Màría
54	R. J. Terrile y P. Laques
55	NASA/JPL
56	A. P. Ingersoll, P. Read, T. Maxworthy, L. G. Redekopp, G. P. Williams y Magda Màría
57	P. S. Marcus, J. Sommeria, NASA/JPL
58-59	Albert E. Zwiazek, Superstock
60	Andrew Christie
61	Fabricación Tri-Estatal (*arriba, izq.*), Zambelli Int. (*arriba, drcha.*), J. Gillmoure (*abajo*)
62	Andrew Christie
63	© Estate of H. E. Edgerton, cortesía de Palm Press, Inc.
64	Museo de Arte Metropolitano (*arriba*), Grucci (*abajo*)
67	Larry Brewer
68	Tom Prentiss
69	G. Kiss (*arriba drcha.*), T. Prentiss (*abajo, drcha. e izq.*)
70	J. R. Newhouse (*arriba, abajo izq.*), G. Kiss (*abajo, drcha.*)
71	Joseph R. Newhouse
74	Johnny Johnson
75	Gabor Kiss
76-79	Johnny Johnson
81	Andrew Christie
82	Edward Bell
83	Russell Ruthen
84	Edward Bell
85-91	PLS, P. Horber



LA FOTOGRAFIA DE LA PORTADA ilustra una sintonía pirotécnica en rojo, blanco y azul. La química ha arrebatado el velo de misterio que rodeaba a la artesanía de los fuegos artificiales al poner de manifiesto en virtud de qué reacciones se producen los colores, las formas y los sonidos característicos (véase "Pirotecnia", de John A. Conkling, en este mismo número).

INVESTIGACION Y CIENCIA

DIRECTOR GENERAL Francisco Gracia Guillén
DIRECTOR EDITORIAL José María Valderas Gallardo
DIRECTORA DE ADMINISTRACIÓN Pilar Bronchal Garfella
PRODUCCIÓN César Redondo Zayas
SECRETARÍA Purificación Mayoral Martínez
EDITA Prensa Científica, S. A. Viladomat, 291 6º 1ª - 08029 Barcelona (ESPAÑA)
Teléfonos 321 81 91 - 321 83 48 Telefax 419 47 82

SCIENTIFIC AMERICAN

EDITOR Jonathan Piel
BOARD OF EDITORS Armand Schwab, Jr., *Managing Editor*; Timothy Appenzeller, Laurie Burnham, *Associate Editors*; Timothy M. Beardsley; Elizabeth Corcoran; John Horgan; June Kinoshita; Philip Morrison, *Book Editor*; John Rennie; Philip E. Ros; Ricki L. Rusting; Russell Ruthen, Paul Wallich; Karen Wright
PUBLISHER John J. Moeling, Jr.
ADVERTISING DIRECTOR Robert F. Gregory
PRESIDENT AND CHIEF EXECUTIVE OFFICER Claus-Gerhard Firchow
CHAIRMAN OF THE BOARD Georg-Dieter von Holtzbrinck
CHAIRMAN EMERITUS Gerard Piel

SUSCRIPCIONES

Prensa Científica S. A.
Viladomat, 291 6º 1ª
08029 Barcelona (España)
Teléfonos 321 81 91 - 321 83 48

Precios de suscripción, en pesetas:

	Un año	Dos años
España	5500	10.000
Extranjero	6200	11.500

Ejemplares sueltos:

Ordinario: 500 pesetas
Extraordinario: 650 pesetas

- Todos los precios indicados incluyen el IVA, cuando es aplicable.
- En Canarias, Ceuta y Melilla los precios incluyen el transporte aéreo.
- El precio de los ejemplares atrasados es el mismo que el de los actuales.

DISTRIBUCION

para España:

MIDESA
Carretera de Irún, km. 13,350
(Variante de Fuencarral)
28049 Madrid Tel. 652 42 00

para los restantes países:

Prensa Científica, S. A. Viladomat, 291 6º 1ª - 08029 Barcelona
Teléfonos 321 81 91 - 321 83 48

ISSN 0210-136X Dep. legal: B. 38.999-76

Fotocomposición: Tecfa. Línea Fotocomposición, S.A. Almogàvers, 189 - 08018 Barcelona
Fotocromos reproducidos por T2, Llull, 145 - 08018 Barcelona
Imprime Rotographik, S.A. Ctra. de Caldes, km 3,7 - Santa Perpètua de la Mogoda (Barcelona)

Printed in Spain - Impreso en España

Hace...

...*cincuenta años*

"Por su propia constitución, el aeroplano será siempre más rápido que el helicóptero; podemos concederle unas 500 millas por hora frente a unas posibles 300 millas por hora del avión de alas giratorias. Sin embargo, en aplicaciones militares defensivas, el helicóptero acomete operaciones que le están vedadas al aeroplano. Citamos a Igor Sikorsky: 'Para interponer una defensa eficaz contra los bombarderos o bombarderos en picado, el helicóptero me parece ideal. Es capaz de permanecer quieto en el aire, proporcionando así un asentamiento de ametralladora estable desde el cual el artillero esperaría el momento en que el bombardero deje de zigzaguear y enfile hacia su presa en línea recta. Una vez lanzado desde un acorazado o un crucero, un avión sólo puede rescatarse posándose en el mar, si éste se halla lo bastante encalmado, e izándolo a bordo. Pero un helicóptero necesitaría únicamente un espacio en cubierta, un cuadrado de unos 12 metros de lado, desde el que despegar y aterrizar.'

"Mientras que gran parte de Europa ha padecido el poder destructor de las bombas en los últimos meses, hay por lo menos un país que ha sabido hacer un uso positivo de tan devastadores artefactos. Suiza los está empleando para hacer caer los aludes de manera controlada. En los Alpes helvéticos se producen todos los años numerosas víctimas a consecuencia de los aludes naturales. Por ello, los suizos han fundado una Estación para el Estudio de los Aludes de Nieve en una localidad próxima a Davos, que es la única de su género en el mundo. La estación puede pronosticar los aludes y lanzar las alarmas correspondientes. Además, si es posible llegar hasta la ubicación del alud, puede provocarse su desplome mediante granadas de mano. Si se trata de aludes de grandes proporciones, su caída se provoca con morteros de trinchera."

"En estos últimos años se ha materializado el sueño centenario de los metalúrgicos de producir lingotes, chapas y bandas en una sola fase a partir del metal fundido. Se fabrican ya lingotes de aleaciones de cobre, aluminio y magnesio y parece ser que a éstos no tardará en sumarse el acero. En ese proceso de colada continua, así se le conoce, el metal fundido fluye desde un horno de mantenimiento hasta las lingoteras, las cuales están equipadas de camisas de enfriamiento; luego, conforme el lingote sale de la lingotera, es enfriado por rociado con agua pulverizada y avanza por gravedad hacia una sierra o soplete que lo corta en trozos de la longitud deseada. Pese a que estos procesos sean muy simples, han debido allanarse muchas dificultades, en razón de los grandes cuidados que deben dispensarse al control de la temperatura y de la velocidad de funcionamiento y a las lingoteras."

...*cien años*

"Hasta hace muy poco, se creyó siempre que el calor desempeñaba un papel primordial en la laminación del acero y que, sin su concurso, no había forma de conseguir las alteraciones precisas para obtener las configuraciones de grano deseadas. Un diario de Chicago acaba de anunciar un cambio en la manipulación industrial de estos productos que echa completamente por tierra las antiguas teorías. De barras de acero frías se sacan alambres por laminación con la misma facilidad que si el metal estuviera caliente; y no sólo eso: el proceso dobla la resistencia a la tracción. Evidentemente, si por laminación en frío es posible obtener alambre a partir de barra con tales resultados, ¿por qué no plancha de acero para barcos u otras aplicaciones? ¿Por qué no también raíles de ferrocarril? A dónde pueda apuntar un hecho de tal novedad se encuentra fuera del alcance de la visión de los más agudos."

"En el levantamiento revolucionario de 1849, una compañía de Honveds, nombre que recibe la milicia húngara, cayó en las vicisitudes de la guerra y sus cuerpos fueron confiados a las aguas del Echoschacht, lago de considerable profundidad no lejos de Hermannstadt. Transcurridos ya cuarenta y un años, los cuerpos han sido devueltos a la luz del día y sometidos a un estudio anatómico, esmerado y minucioso. El doctor Koenig, médico forense de Hermannstadt, los halló en perfecto estado de conservación, al igual que muchos de los preparados de los museos anatómicos. Las epidermis presentaban un color gris blanquecino y los músculos, bermellón, ofreciendo un tacto similar al de la carne fresca de una tienda de carnicería."

"¿Es que los miembros de la gran profesión médica, quienes viven para frenar el dolor, la enfermedad, la muerte,..., deben prestarse, cualesquiera que sean las circunstancias, a desempeñar el repugnante papel de verdugos públicos? Al menos por lo que a mí respecta, rotundamente *no*. Por el contrario, vaya mi protesta más enérgica contra cualquier procedimiento de ejecución, como la electricidad u otro, cuya práctica reclame necesariamente la intervención del hombre de ciencia. Ya es hora de que nuestros médicos carcelarios se pregunten si deben seguir condescendiendo a cumplir el tristemente llamado deber de atestiguar profesionalmente un asesinato capital. ¡Pero asumir un papel destacado, *nunca*!"

"Cualquiera que sea la opinión que se sustente acerca de las carreras de caballos y su repercusión moral, no deja de ser cierto que ejercen una atracción irresistible en un gran sector del público. El hipódromo eléctrico ocupa un lugar equidistante de las carreras reales y las de juguete. Esta instalación realizada en Niza (Francia) comprende un motor de gas de 12 caballos de potencia que acciona una dinamo Rechniewsky de devanado doble, la cual envía corriente a seis motores eléctricos de una potencia de 1000 watt cada uno; ponen en movimiento a un caballo con ruedas, siendo la velocidad, el arranque y la parada de cada una de éstas completamente independientes de las de las demás."

El gran debate sobre el clima

El efecto de invernadero y la posibilidad de un calentamiento global constituyen objeto de discusión científica y política. ¿Debemos tomar medidas ahora para evitar consecuencias que no podemos prever?

Robert M. White

En las postrimerías del siglo décimo, millones de personas se prepararon para el apocalipsis, creyendo que el inminente año 1000 era el mismísimo milenio, el fin de los cielos y la Tierra profetizado en el último libro del Nuevo Testamento. No debiera sorprendernos que la perspectiva del cercano Día de la Ira aterrorizase a personas normalmente sensatas, incitándoles a obrar de manera precipitada y, visto en retrospectiva, necia. Algunos se desprendieron de todas sus posesiones; otros se entregaron a duras penitencias que expiaran sus malas acciones.

En este decenio final del siglo xx, otra clase de apocalipsis causa general preocupación. Esta vez, la mano de Dios ha sido sustituida por agentes más visibles: chimeneas que arrojan humo, automóviles que se mueven con gasolina, plantas generadoras de energía y la voraz destrucción de los bosques, pueden estar caldeando progresivamente un ambiente sobrecargado. El calentamiento climático global, mantienen algunos, amenaza incluso la habitabilidad del planeta. Otros sostienen que las predicciones de un colapso del medio ambiente carecen de suficiente fundamento y nos están forzando a actuaciones políticas irreflexivas. ¿Es nuestro planeta "La Tierra en peligro", según pretendía la revista *Time* en su artículo de fondo del final del año 1988? ¿O se trata, como decía la revista *Forbes*, de "El pánico del calentamiento global: un caso de reacción exagerada"?

El debate que ha llegado a los medios de comunicación refleja la incertidumbre reinante entre climatólogos y geofísicos. Algunas de las autoridades mundiales sobre la atmósfera perdieron, hace poco, la compostura académica en las páginas de *Science*. Las acusaciones de "ciencia barata" y "ciencia por consenso" manifiestan el tono agrio del debate en el seno de la comunidad científica. Miembros de la Academia Nacional de Ciencias de los EE.UU. (NAS), entre ellos uno de sus antiguos presidentes, sostienen que se está induciendo a los responsables políticos a tomar medidas imprudentes que se basan en datos de escasa fiabilidad. A este punto de vista se opone la reciente declaración de la Unión de Científicos Preocupados, que exige la intervención inmediata del gobierno. Fue firmada por 52 premios Nobel y más de 700 socios de la mencionada institución.

Pese a la falta de certezas científicas, hay grupos públicos y privados empeñados en ir más lejos que nadie a la hora de reclamar acciones drásticas e inmediatas para "estabilizar" el clima global. Desde Washington hasta Toronto y La Haya, desde El Cairo hasta Moscú, se convocan conferencias internacionales de expertos y líderes políticos que exigen acción. Mikhail S. Gorbachev, George Bush, Margaret Thatcher y François Mitterand comparten puntos de vista similares sobre el problema del calentamiento global.

En los Estados Unidos asistimos hoy a un intenso debate en el seno de la administración Bush sobre el modo de actuar del gobierno. Entre la espada de la llamada a la acción por parte de la Agencia de Protección Ambiental (EPA) y el Congreso y la pared de las llamadas a la prudencia de su asesor científico y de su secretario general, Bush convocó dos conferencias importantes para abordar el problema del calentamiento global. La primera, celebrada en abril, reunió a los directores de las agencias científicas, económicas y ambientales de muchos gobiernos. La segunda será una primera reunión de gobiernos, fijada para comienzos de 1991, que iniciará la negociación de un acuerdo internacional para estabilizar el clima global.

Mientras la Casa Blanca se debate en la duda, el Congreso ha mostrado una inusitada preocupación por los asuntos ambientales. Se han introducido muchas propuestas legislativas sobre el calentamiento general. A la cabeza del esfuerzo conjunto de ambos partidos han estado el senador Timothy E. Wirth, de Colorado, el senador Al Gore, de Tennessee, y la diputada Claudine Schneider, de Rhode Island. La normativa es, en buena parte, amplia y de largo alcance. Abre vías para intervenir en los frentes de la energía, la agricultura y el transporte, así como para intensificar la investigación.

Las acciones propuestas modificarían, de raíz, las funciones vitales de la economía humana. Abarcarían medidas muy dispares: obtener mayor rendimiento de la energía, sustituir la mezcla de petróleo y carbón por gas natural, apoyarse más en las fuentes de energía renovables y usar más energía nuclear y solar. Entre las medidas podría también incluirse el fomentar la reforestación, la gradual eliminación del uso de halocarburos (CFC) y la modificación de ciertas prácticas agrícolas.

Estas iniciativas alterarían la técnica y la economía energéticas. Habría de cambiar, asimismo, nuestro empleo del suelo y el agua. Podría frenarse el desarrollo económico de

ROBERT M. WHITE es presidente de la Academia Nacional de Ingeniería de los Estados Unidos de América. Ha sido jefe de la Oficina sobre el Clima y secretario de la Administración Nacional Oceánica y Atmosférica. Ha estado también al frente de la Corporación Universitaria para la Investigación Atmosférica y de la Sociedad Meteorológica Americana. En 1979 presidió la primera Conferencia Mundial del Clima de la Organización Meteorológica Mundial.

las naciones que pende de combustibles fósiles. Se pondría el acento en el problema que plantea el crecimiento de la población. Pero, ¿cómo pueden redactarse con tanta rapidez propuestas de política nacional e internacional para hacer frente al espectro del cambio climático, cuando falta el acuerdo sobre sus fundamentos científicos y apenas se han estimado sus costes económicos y sociales? En la raíz de esta manera de pensar descubrimos la confluencia de diversas fuerzas científicas, económicas y ambientales.

La idea de que la acción del hombre podría cambiar la composición de la atmósfera y, con ello, el clima mundial tiene una larga historia. Ya en el decenio de 1860 se lanzó la idea según la cual bastaban ligeros cambios en la composición atmosférica para producir importantes variaciones en el clima. Los aumentos de concentración del dióxido de carbono (CO_2) y otros gases traza en la atmósfera podían contribuir al calentamiento de invernadero, llamado así porque esos compuestos permiten que la energía solar llegue a la superficie de la Tierra, calentándola, e impiden que gran parte de la energía reemitida se pierda en el espacio.

A fines del siglo XIX, Svante A. Arrhenius, científico sueco, calculó cómo afectaban los cambios del contenido de dióxido de carbono a la temperatura de la superficie de la Tierra. Estimó que la duplicación del dióxido de carbono produciría un calentamiento global de unos cuatro a seis grados Celsius, no muy lejos de los cálculos modernos. No obstante, hubo que esperar a la puesta en marcha del Año Geofísico Internacional, un experimento a escala mundial realizado en 1957 y destinado a registrar el entorno global, para disponer de datos científicos que confirmasen el aumento del dióxido de carbono en la atmósfera.

Roger Revelle, entonces director de la Institución Scripps de Oceanografía, su colega Hans E. Suess y su alumno C. David Keeling emprendieron tales mediciones. Revelle había sostenido durante mucho tiempo que los humanos estaban realizando, sin darse cuenta, un experimento geofísico con la atmósfera al quemar carburantes fósiles. Decidido a vigilar su contenido de dióxido de carbono, persuadió a Keeling para que desarrollara los instrumentos necesarios.

Los aparatos de medición se colocaron en el observatorio meteorológico de Mauna Loa, en Hawai, a unos 3300 metros de altitud. Desde 1957, los datos registrados pusieron de manifiesto un incremento sistemático de la cantidad de dióxido de carbono. Las observaciones de Keeling se comprobaron en el polo Sur y en otros lugares del mundo entero. Hasta ahora, el cambio producido, desde 290 partes por millón en 1880 hasta 352 en 1989, representa un aumento de más del 20 % en ese siglo transcurrido.

Poco antes del comienzo de Año Geofísico Internacional, en el extremo opuesto de los Estados Unidos estaba produciéndose otro adelanto clave para desentrañar las consecuencias climáticas del aumento de las emisiones de dióxido de carbono. En el Instituto de Estudios Avanzados de Princeton, Nueva Jersey, y bajo la dirección de John von Neumann, matemático de prestigio mundial, se realizaban los primeros ensayos de representación matemática de la atmósfera con un ordenador digital.



1. BARCAZAS VARADAS por el descenso del nivel de agua, el más bajo conocido en el río Mississippi. Debióse a la sequía de 1988, la más grave registrada. Aquel caluroso verano motivó que la gente comenzara a pensar en el efecto de invernadero y se preocupara por la política a adoptar.

El grupo de científicos de von Neumann, jóvenes y brillantes, lo encabezaba Jule G. Charney. Quien iba a ser reconocido padre de la predicción numérica del tiempo, y con credenciales para ser considerado la figura norteamericana más importante en la transformación de la predicción meteorológica de arte en ciencia, Charney mostró la viabilidad del uso de los ordenadores para acometer esa tarea. Von Neumann y Charney calcularon la primera predicción para un período de 24 horas en 1950, con un primitivo ordenador digital, el ENIAC, que el Cuerpo de Transmisiones del Ejército de los EE.UU. tenía en Nueva Jersey.

Oteando más allá de estos esfuerzos, von Neumann llamó, a la del clima, la "predicción infinita". Uno de los jóvenes investigadores del grupo de Princeton, Norman Phillips, realizó el primer ensayo de modelo de atmósfera global en 1956. Fue coincidencia el que más tarde, en 1963, la Administración Nacional Oceánica y Atmosférica (NOAA) estableciera un laboratorio insólito en el recinto de la Universidad de Princeton bajo la dirección de Joseph Smagorinsky, joven y exigente investigador de férrea voluntad que había formado parte del grupo de von Neumann. El laboratorio se dedicaba por completo al modelado matemático de la atmósfera usando los mayores y más rápidos ordenadores disponibles.

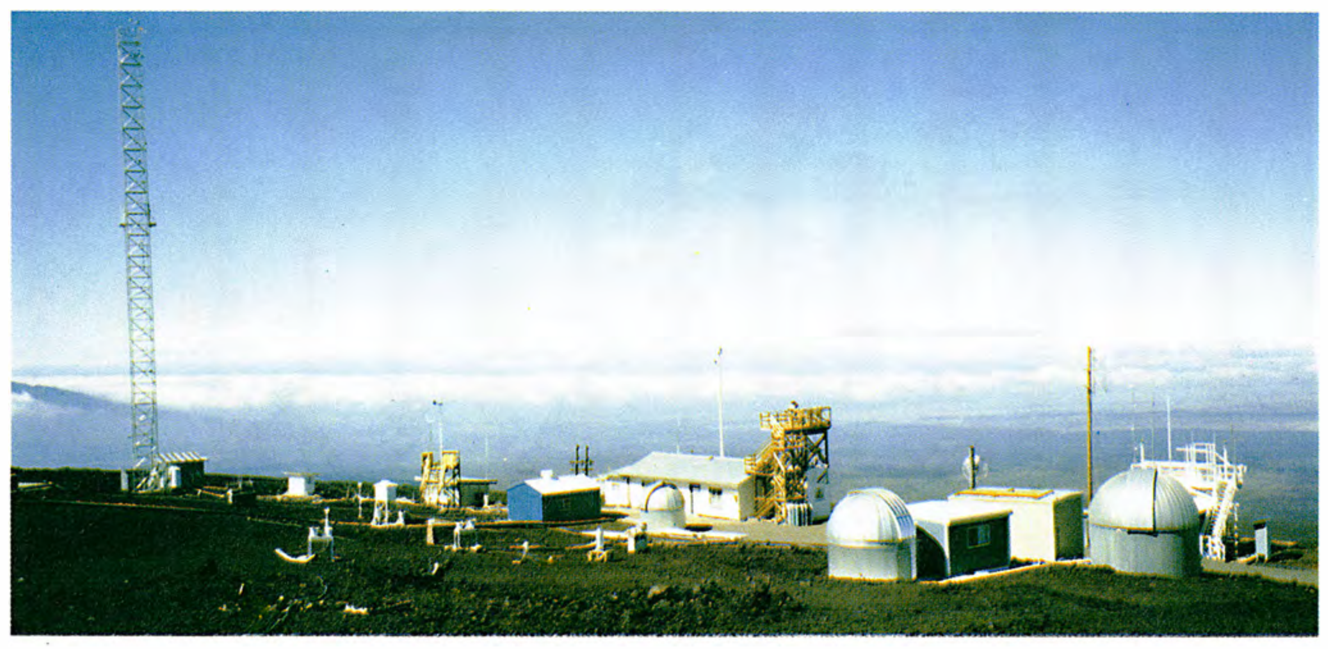
El Laboratorio de Dinámica de Fluidos Geofísicos, así se bautizó el centro, acogió investigadores de distintas naciones interesados en este nuevo enfoque del estudio de la atmósfera. Entre ellos se encontraba Syukuro Manabe. Este joven japonés, modesto y retraído, pero dedicado en cuerpo y alma a su trabajo, desarrolló el primer modelo climático en colaboración con su colega Richard T. Wetherald en los años sesenta. En 1975 calcularon que la duplicación del contenido de dióxido de carbono en la atmósfera produciría un calentamiento climático global de unos tres grados Celsius, en promedio, sobre toda la superficie de la Tierra. Este cálculo, corroborado luego en muchos laboratorios, no ha sufrido variaciones sustanciales.

Las observaciones de Keeling, sumadas a los cálculos de Manabe y Wetherald, provocaron la oleada de investigaciones sobre el cambio climático que ha marcado los dos últimos decenios. De entonces acá se han multiplicado los estudios en muchas partes del mundo, entre ellas Europa y la Unión Soviética. En los EE.UU., el Consejo Nacional de Investigación acometió varios proyectos en 1966, 1977, 1979, 1983 y 1987. Estas investigaciones estuvieron dirigidas por expertos autorizados: Gordon J. F. McDonald, Revelle, Thomas F. Malone, Charney, William A. Nierenberg y el economista William D. Nordhaus.

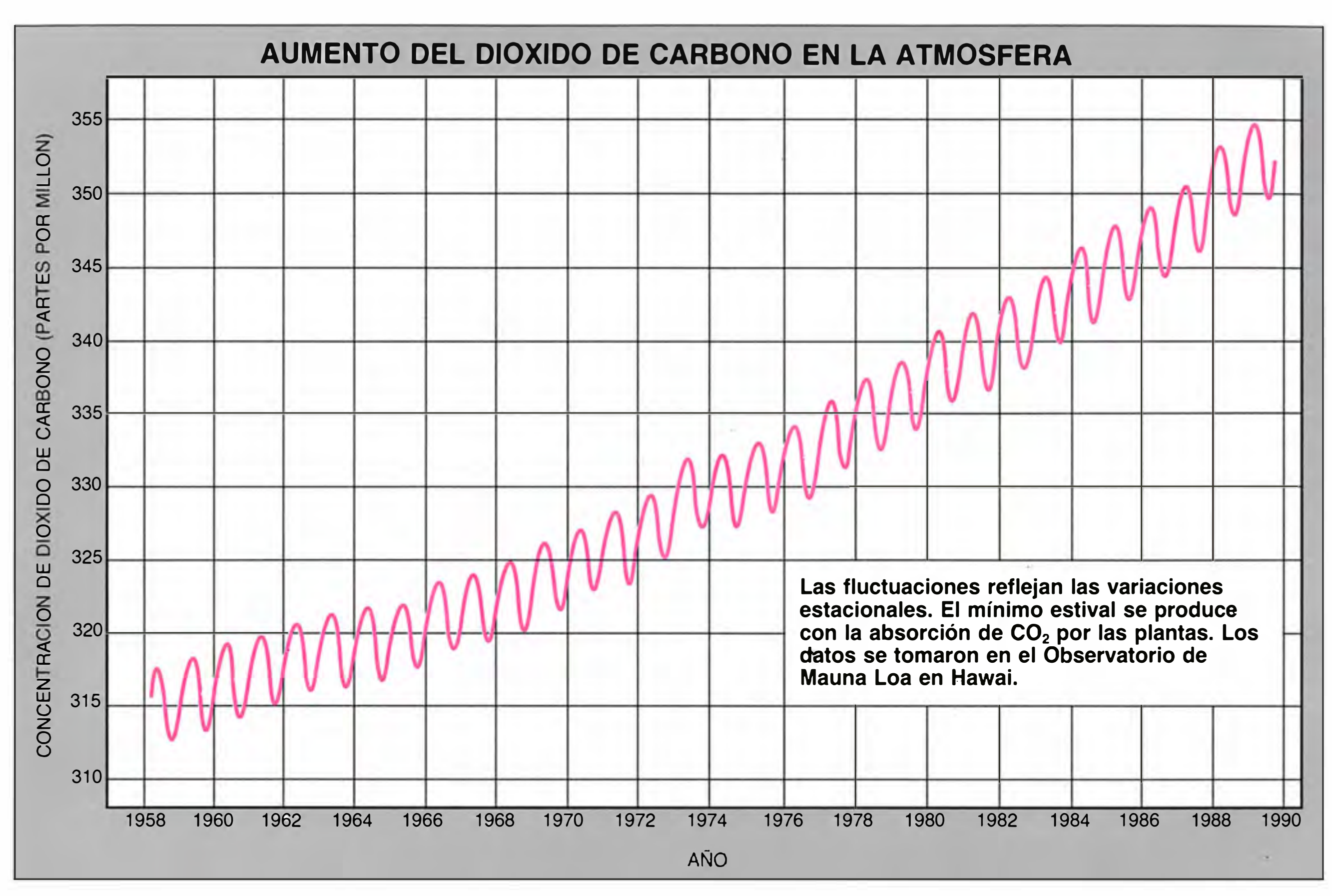
Ahora bien, como no había consecuencias inmediatas para la salud humana ni manifestación tangible de cambio climático, estos trabajos tardaron en despertar la preocupación política. El estudio que más influyó, en el ámbito de la política fue el preparado en 1979 a petición de Frank Press, actual presidente de la NAS, que era entonces el asesor científico del presidente Jimmy Carter. También en 1979, la Organización Meteorológica Mundial, con sede en Ginebra, consciente de la potencial importancia global del problema, convocó la primera Conferencia Mundial del Clima.

Poco a poco, la toma de conciencia científica de que la humanidad podría estar provocando un desarreglo de proporciones planetarias comenzó a hacer mella en el mundo político. Aunque se debatió largo y tendido sobre la validez de las predicciones de los modelos de ordenador, las observaciones del aumento de los gases de invernadero eran, sin embargo, precisas, bien medidas y comprobadas en muchas partes del mundo. Las corroboraron datos adicionales que evidenciaban aumentos en otros gases de invernadero, como el metano (gas natural) y los CFC.

Mientras tanto, en los EE.UU. se habían creado nuevos grupos de modelado matemático. Al NOAA, había que agregar los de la Administración Nacional de Aeronáutica y del Espacio (NASA), el Departamento de Energía y la Fundación Nacional de Ciencia (NSF). Los directores de esos laboratorios se convirtieron en "gurúes" del cambio climático. Incisivos y originales en su trabajo, Stephen H. Schneider, del Centro Nacional de Investigación Atmosférica de la NSF en Boulder, y James E. Hansen, director del Instituto Goddard de Estudios Es-



2. OBSERVATORIO METEOROLOGICO DE MAUNA LOA, Hawai. Está situado a unos 3300 metros. Los datos tomados, que se inician en 1957 y siguen hasta el presente, suministraron los primeros registros del continuo aumento de los niveles de dióxido de carbono en la atmósfera.



paciales de la NASA, pasaron a ser expertos asiduos en las innumerables sesiones que al tema consagraron los comités del Congreso.

Aunque los modelos matemáticos de todos los grupos daban resultados similares, diferían en los detalles de la distribución geográfica de los cambios climáticos. Todos pronosticaban que un aumento del dióxido de carbono conduciría a un calentamiento gradual, pero el calendario de este calentamiento dependería de la tasa de uso global de la energía. Todos concordaban en que las hipótesis razonables sobre el futuro consumo global de energía abonaban, sin excepción, el pronóstico según el cual, hacia la mitad del siglo próximo, se habría duplicado la concentración de dióxido de carbono.

Ahora bien, el aumento de temperatura que esta duplicación del dióxido de carbono produciría variaba bastante de un modelo a otro. Algunos predecían aumentos pequeños, del orden de un grado Celsius; otros lo cifraban incluso en cinco grados Celsius. La diferente cuantía se convirtió en elemento central del debate en torno a la garantía que los modelos ofrecían para justificar acciones políticas. No era cuestión menor, ni muchísimo menos, que el incremento real estuviera en uno u otro extremo de ese margen. En el extremo inferior, la capacidad de adaptación normal de la sociedad bastaría, quizás, para acomodarse al nuevo clima. Los cambios hacia el extremo superior llevarían consigo graves desarreglos.

Estos cambios de temperatura previstos pueden parecer inocuos, puesto que en el curso diurno y estacional se experimentan variaciones de esa magnitud. Pero pueden percibirse todas sus consecuencias si se considera que sólo se necesitó una disminución de un grado Celsius en la temperatura media de Europa para provocar la Pequeña Edad del Hielo, sucesión de gélidos siglos que duró, aproximadamente, desde 1400 hasta 1800. Se cree que la diferencia de temperatura entre el final del último gran periodo glacial, hace 12.000 años, y el presente es de cinco grados Celsius. Además, las previsiones indican que el hemisferio Norte experimentaría, en apenas medio siglo, un cambio de temperatura sin precedentes, entre 10 y 50 veces más rápido que el cambio ocurrido desde el último periodo glacial.

A quienes no están familiarizados con los modelos matemáticos ni con el uso de ordenadores en tales previsiones se les puede perdonar que se sientan confusos, y hasta irritados, por las discrepancias de los resultados. Entre investigadores se sobreentiende que los modelos matemáticos son sólo aproximaciones que intentan simular los procesos que rigen el comportamiento atmosférico. La complejidad de la atmósfera es tal, que resulta imposible representarla con mucho detalle en esos modelos matemáticos. Hay que limitarse a determinadas características y establecer hipótesis adicionales sobre cómo interactúan la atmósfera y los océanos, cómo varía el ritmo de absorción de dióxido de carbono por los océanos y cómo afectan las nubes al intercambio de energía entre la Tierra y la atmósfera. Ni siquiera los mayores ordenadores pueden representar la atmósfera, los océanos y la superficie terrestre con la deseada minuciosidad. La verdad es que los investigadores determinan las condiciones de la atmósfera abordándola a través de observaciones separadas entre sí por unos 500 kilómetros.

La demanda de la intervención política se está realizando con ese telón de fondo de incertidumbre. Por un lado se argumenta que, de haber alguna probabilidad de certeza en las predicciones de los modelos, las consecuencias serían calamitosas y se haría imperiosa la acción inmediata para detener el cambio climático. El punto

de vista opuesto, no menos sensato, objeta la irracionalidad de emprender acciones de enormes consecuencias económicas y sociales ante nuestra inseguridad en lo científico y nuestro desconocimiento de los costes económicos. John H. Sununu, secretario general de la Casa Blanca, se convirtió en portavoz de esta postura en declaraciones hechas con motivo de la reunión anual de la Academia Nacional de Ingeniería en el otoño de 1989:

"Aunque comparto la idea de que [el calentamiento global] constituye un asunto crítico, el hecho es que los modelos con los que se realiza el análisis y con los que se configura la política, aun siendo buenos, se siguen fundando en elementos cuya longitud y anchura es de cientos de kilómetros y su espesor no supera decenas de kilómetros. Sospecho que ningún experto en simulaciones de ingeniería se encontraría cómodo si tuviese que tomar decisiones importantes cuando los elementos trascendieran, en varios órdenes de magnitud, los detalles de los que se obtuviese la información. Y lo cierto es que caminamos hacia una política internacional vinculante basada en conclusiones sacadas por políticos que no perciben en absoluto la diferencia existente entre los niveles de confianza que deberían tener y los niveles de confianza que desean tener. Un sistema no es válido sólo porque le dé a uno las respuestas que desee. Pese a ello, buena parte de la política se está configurando a tenor de ese principio."

La solución del dilema debería ser sencilla: puesto que la concentración de dióxido de carbono en la atmósfera ha aumentado en más del 20 por ciento en los últimos cien años, tendríamos que poder comprobar el calentamiento con los registros de la temperatura media global durante el mismo período. Eso es lo que han tratado de hacer los investigadores, pero resulta mucho más difícil de lo que parece. El problema estriba en que el clima se halla siempre en un estado de fluctuación natural. Separar, de los cambios naturales, los causados por el aumento del dióxido de carbono no es un proceso fácil. Además, el registro de temperaturas se ha creado con observaciones dispersas e irregulares, que no se tomaron específicamente para determinar las condiciones climáticas.

Ello no obstante, y a través de análisis minuciosos de estos archivos de temperaturas realizados por investigadores de los EE.UU. y el Reino Unido, se ha tratado de descubrir si se ha producido un calentamiento global y si éste concuerda con las predicciones de los modelos. El punto de vista dominante es que los datos climáticos recogidos a lo largo de los últimos cien años para el globo entero revelan un aumento neto de temperatura que varía desde 0,3 hasta 0,8 grados Celsius. Conclusión, sin embargo, a la que hay que poner sordina con otro resultado: no se han detectado aumentos de temperatura similares, en ese lapso temporal, en los EE.UU., donde las observaciones son numerosas y fidedignas.

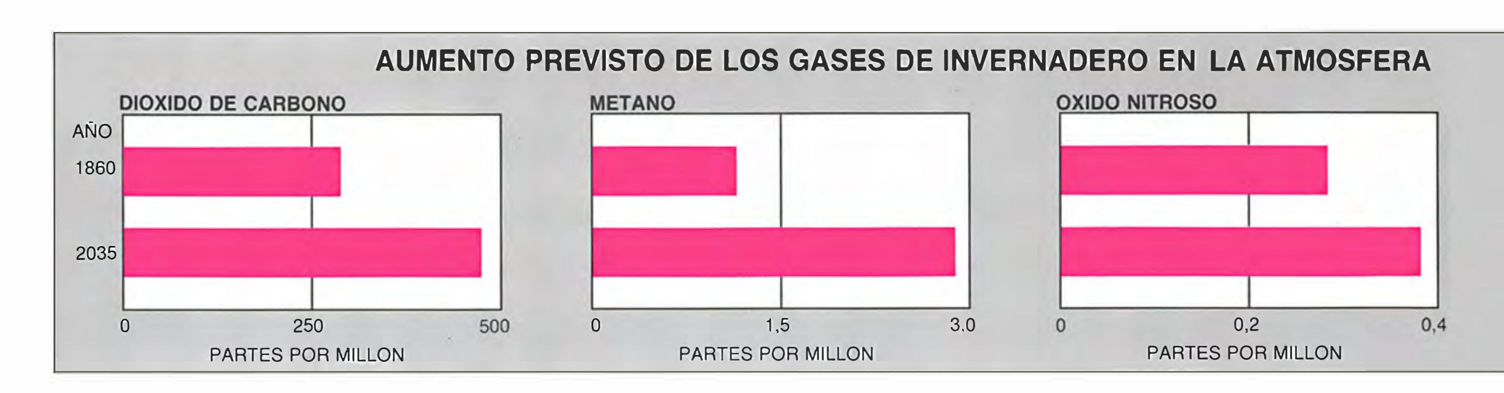
Aun cuando aceptemos la realidad de ese aumento, queda un interrogante que los investigadores no han sabido despejar; nos referimos a la naturaleza de la subida global de temperatura: ¿es una fluctuación natural u obedece a una mayor concentración de gases de invernadero? Lo único que puede afirmarse es que el aumento observado guarda coherencia con la cifra menor de la escala de aumentos de temperatura predichos por los modelos informáticos. En consecuencia, los registros de temperatura y los modelos de ordenador respaldan a quienes creen que los datos justifican la intervención inmediata y a quienes piensan que las pruebas son todavía insuficientes.

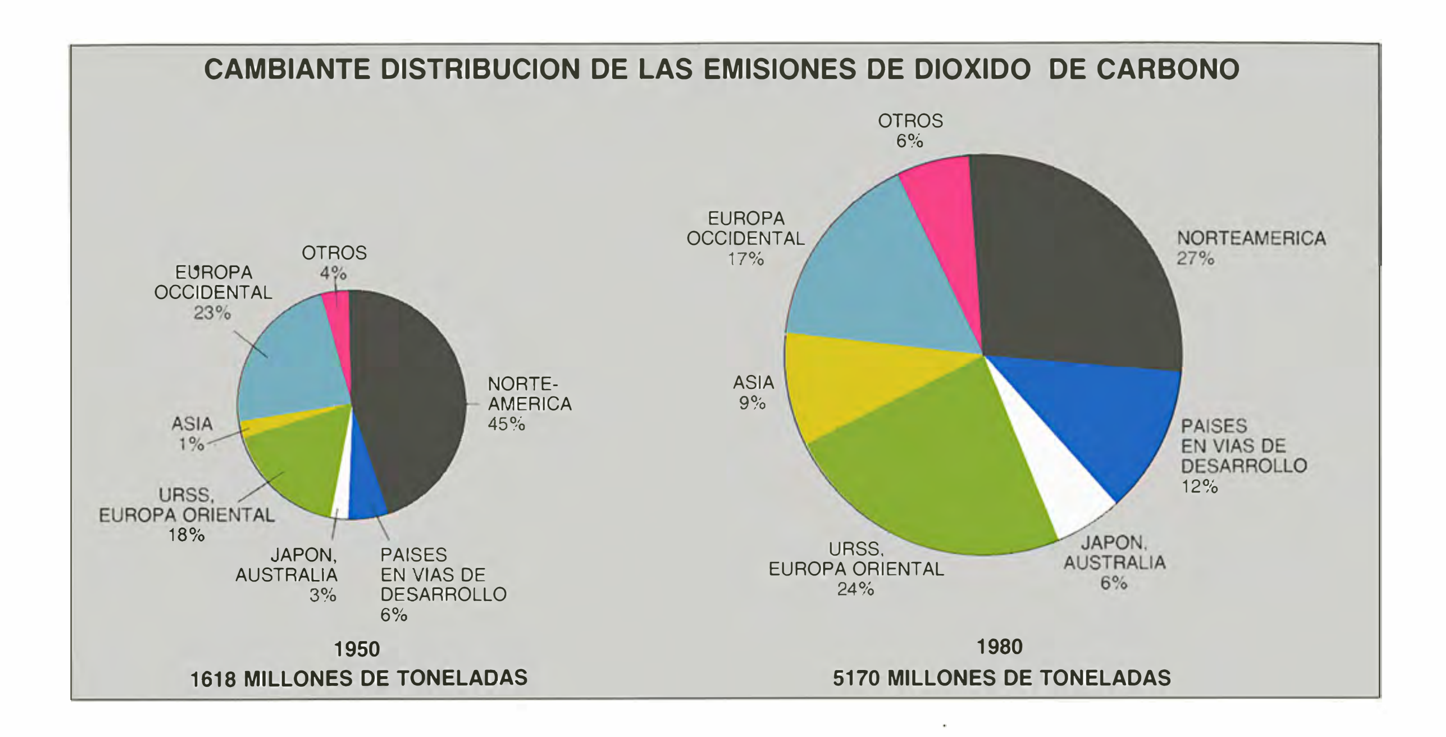
En mi opinión, el apremio de una actuación política inmediata fue catalizada por la desastrosa sequía padecida en Estados Unidos durante el verano de 1988. Por culpa de ella, una de las peores de que se tenga noticia, el río Mississippi bajó tanto que la navegación se hizo imposible en largos trechos, se vio amenazado el abastecimiento urbano de agua y se perdieron las cosechas del cinturón granero. La administración y los ciudadanos se preguntaron si no era aquélla una manifestación del efecto de invernadero. De hecho, los registros ponen de relieve que, en los EE.UU., cinco de los años del decenio de 1980 figuran entre los más cálidos de que se tengan datos, y la temperatura media de esa década fue la más alta desde que comenzaron los registros con instrumentos.

Espoleadas por el calor y la sequía, las comisiones del Congreso se preocuparon de averiguar si había llegado el efecto de invernadero. Las comparecencias de los expertos no ofrecieron nada de particular, salvo una declaración de Hansen. Aseveró contar con un 99 por ciento de seguridad de que el efecto de invernadero había llegado; lo probaba la sucesión de años cálidos del decenio de 1980. El público tomó nota entonces de ello. La opinión de Hansen movió a los miembros del Congreso a considerar que quizá lo prudente sería promulgar con rapidez leyes encaminadas a proteger, de consecuencias catastróficas, la habitabilidad del planeta.

Se sucedieron las comparecencias. Expertos en el comportamiento de la atmósfera y demás ciencias relacionadas con el entorno, así como los ecologistas, empezaron a tomar partido en pro o en contra de una intervención política inmediata. La reacción de los ecologistas fue rápida y vocinglera. Varios grupos ecologistas y científicos pidieron acuerdos internacionales que restringiesen las emisiones de gases de invernadero.

Llegados a este punto, algunos investigadores influyentes, que creían que las actuaciones políticas empe-





zaban a ir más allá de las pruebas científicas disponibles, hicieron valer sus puntos de vista. Richard S. Lindzen, del Instituto de Tecnología de Massachusetts, y Jerome Namias, de Scripps, el experto estadounidense más cualificado en predicción a largo plazo, escribieron una carta al presidente Bush urgiéndole a no tomar medidas. Otros tres miembros de la NAS, entre ellos su anterior presidente, Frederick Seitz, elaboraron un informe conjunto, publicado bajo los auspicios del Instituto Marshall, en el que se valoraban las bases científicas de tales actuaciones políticas. Recomendaron un intenso programa de investigación sobre la creación de modelos matemáticos. Podía haber explicaciones distintas, señalaban, para el calentamiento observado. El gran debate sobre el clima se había desencadenado.

Los meteorólogos no vieron con buenos ojos la perspectiva de una nueva disputa pública en la que se involucrara su especialidad. Se habían equivocado muchas veces en el pasado. Ya en 1924, Sir Gilbert Walker, entonces jefe del servicio meteorológico de la India bajo el gobierno británico, había detectado relaciones, muy estrechas e insólitas, entre lluvia, temperatura y presiones en el Océano Indico y en el subcontinente indio. Se lanzaron las campanas al vuelo porque, creíase, se había resuelto el problema de la predicción del monzón indio. De haber sido verdad, habría constituido una valiosísima ayuda para la agricultura indostánica. No tardó en comprobarse que las correlaciones poseían escasa consistencia predictiva.

Más tarde, en los decenios de 1940 y 1950, corrió la voz de que la siembra de nubes con hielo seco o cristales de ioduro de plata incrementaba la pluviosidad; basábase en los trabajos de Irving Langmuir, que fue premio Nobel y trabajó en la General Electric Company, y de Vincent J. Schaefer, de la Universidad del estado de Nueva York en Albany. Decenios de investigación sobre las posibilidades de aumentar la cantidad de lluvia, alterar la intensidad de los huracanes tropicales o modificar las granizadas mediante técnicas de siembra de nubes terminaron siempre en fracaso.

A principios de los años ochenta, se avanzó que el polvo lanzado a la atmósfera como consecuencia de un intercambio de proyectiles nucleares entre la URSS y los EE.UU. produciría un "invierno nuclear". La idea se vino abajo cuando Schneider y su colega Stanley L. Thompson mostraron, con los mismos modelos usados en la predicción del clima, que "la hipótesis inicial de un invierno nuclear puede relegarse ahora a un nivel de probabilidad inapreciable".

Ante historia tan poco afortunada, no es de extrañar que muchos meteorólogos se manifiesten reticentes a tomar costosas medidas cuya justificación se apoye en las predicciones del calentamiento global. Pero en ese apremio por intervenir hacen acto de presencia otros factores. El calentamiento global une también a los que están preocupados por la diversidad biológica y la extinción de especies, por el desarrollo económico, el crecimiento demográfico, la contaminación de las ciudades, la precipitación ácida y el empobrecimiento de la capa de ozono.

Los dirigentes políticos, atentos a la sensibilidad del público por la degradación del medio, ven, en estas cuestiones, importantes plataformas y trampolines de medro personal. Los interesados en robustecer la competitividad de la industria norteamericana consideran el uso más eficaz de la energía un paso decisivo en ese camino. Idea de la que participan quienes están preocupados por la dependencia estadounidense de fuentes de energía extranjeras. Entra asimismo en juego la cuestión de la energía nuclear. Por ser los combustibles fósiles fuente principal de dióxido de carbono atmosférico, las estrategias para estabilizar el clima deben contemplar

HALOCARBUROS
0 0,025 0,005
PARTES POR MILLON

fuentes de energía distintas. He aquí, por fin, una justificación que los defensores de la energía nuclear pueden esgrimir con fuerza para apoyar la expansión de centrales por todo el mundo.

El problema del calentamiento global ofrece también una oportunidad para impulsar el "nuevo orden económico" que hace tiempo proponen los países del Tercer Mundo. La actuación internacional requerirá que se preste a esas naciones ayuda tecnológica y económica, si han de participar en un esfuerzo global para reducir la contaminación atmosférica o detener la desforestación. En efecto, el mundo se enfrenta a la posibilidad de que los mayores incrementos en las emisiones de dióxido de carbono se produzcan en los países en vías de desarrollo, conforme su necesidad de crecimiento vaya acompañada por una mayor demanda de energía.

De igual manera, los interesados en detener el crecimiento de la población, sobre todo en el llamado Tercer Mundo, señalan la probable irresolución del problema del calentamiento global mientras continúe aumentando el número de seres humanos. La gente es, en definitiva, la que consume recursos naturales y energía y la que cultiva la tierra. Sin control de la población, las perspectivas de estabilizar el clima y frenar el empeoramiento de la habitabilidad del planeta son remotísimas.

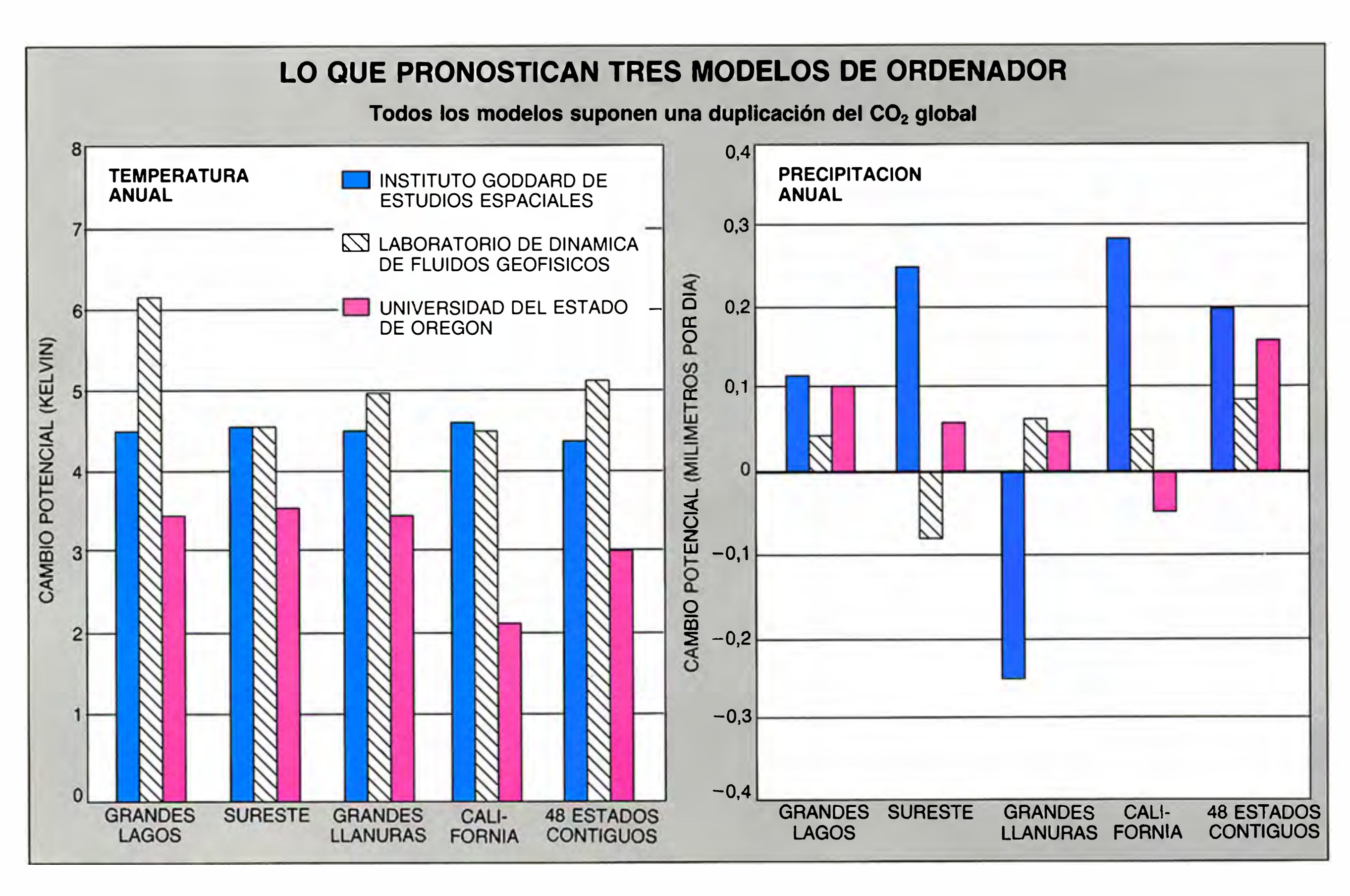
¿Hemos de temer las consecuencias del calentamiento global? Las personas sufren temperaturas extremas en el curso normal de los acontecimientos. La verdad es que no sabemos lo suficiente para predecir la gravedad de las consecuencias. Como el calentamiento no sería uniforme en toda la superficie de la Tierra, habría, probablemente, ganadores y perdedores entre las regiones y las naciones. Unas partes ganarían en calor, otras se harían más lluviosas y unas terceras más secas. Con los datos de que disponemos no cabe predecir quién se beneficiaría y quién saldría perdiendo en semejante redistribución global de los recursos climáticos.

Habría aspectos del calentamiento global muy beneficiosos desde el punto de vista de la agronomía. El incremento del dióxido de carbono fomentaría una fotosíntesis más activa y un crecimiento más rápido de las cosechas, por no hablar de una menor exigencia hídrica por parte de las plantas en una atmósfera enriquecida en CO_2. Como ha dicho Jerry D. Mahlman, director del Laboratorio de Dinámica de Fluidos Geofísicos: "Los políticos no están interesados en lo que nosotros podemos afirmar con garantía. Y sobre las cosas que a ellos les interesan no podemos pronunciarnos con seguridad".

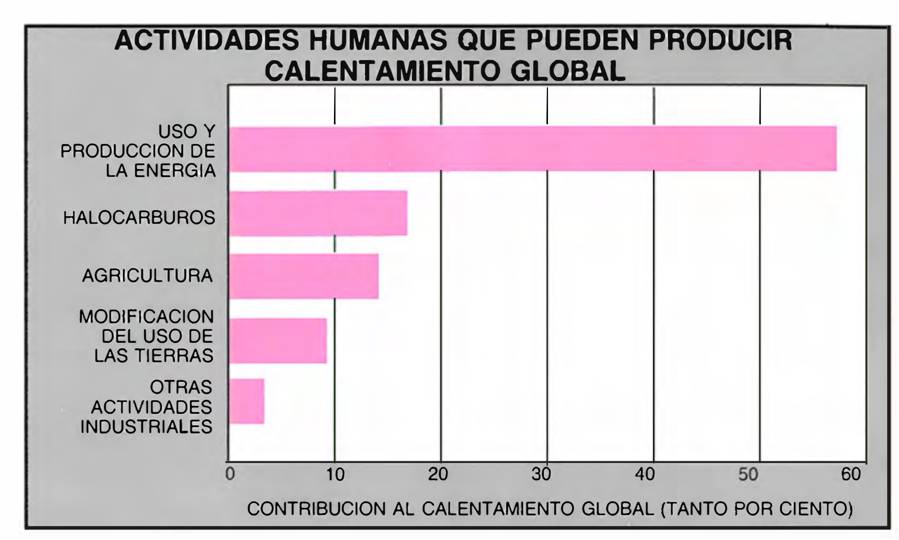
Los modelos, empero, concuerdan en que las regiones polares experimentarán mayor subida de temperatura que los trópicos. Algunas predicciones sobre la elevación de la temperatura en los polos manejan cifras alarmantes: de hasta 10 grados Celsius en promedio en el hemisferio Norte, con sólo leves aumentos en los trópicos.

¿Qué consecuencias generales implican tales cambios en la diferencia térmica entre las regiones polares y las ecuatoriales? Todos los años se presentan diferencias parecidas en el transcurso de las estaciones. En verano, cuando las temperatuas árticas son bastante altas, no estamos sometidos a las grandes borrascas del invierno; los cinturones de precipitación se corren hacia el norte. La parte suroeste de los EE.UU., por citar una región, es entonces muy seca.

Si las regiones árticas experimentaran calentamientos apreciablemente mayores que las ecuatoriales y si los cinturones de precipitación se trasladaran hacia el norte, los países de las zonas templadas y polares septentrionales se beneficiarían bastante, a buen seguro. Su época de crecimiento



se alargaría y se limitaría su volumen de precipitación. Con suelos adecuados, se asentaría la agricultura. Pero esto pertenece al terreno de la especulación.



Lucubraciones que se formulan en forma de "supuestos" a través de la expresión "¿Y si ...?" Es infinito, por desgracia, el número de preguntas condicionales. ¿Y si las corrientes de los ríos del suroeste norteamericano, aprovechados hoy hasta la última gota, se redujeran en un 20 por ciento? ¿Y si subiese la temperatura en el cinturón cerealista y se corriera más hacia el norte la precipitación? ¿Qué pasaría entonces con la agricultura norteamericana, con los recursos disponibles, con la generación de energía, con los parques nacionales y la conservación de la naturaleza?

De las consecuencias que arrastran los cambios climáticos nos hablan documentos históricos que se remontan a miles de años atrás; las hemos visto también en sucesos climáticos recientes. Sabemos, por ejemplo, que los daneses se establecieron en Groenlandia y los vikingos navegaron por el Atlántico norte hasta el continente americano durante un período de clima cálido, alrededor del año 1000. Después, se produjo una importante modificación del clima que mantuvo confinados a los poblados daneses, impidió la ulterior exploración de Norteamérica y anunció el comienzo de la Pequeña Edad del Hielo. En los últimos años hemos presenciado los efectos de la sequía en la región del Sahel, en Africa y en el nordeste del Brasil, así como en algunas partes de Norteamérica.

Tales supuestos sugieren posibilidades apocalípticas. En un reciente episodio de una serie de televisión aparecía el Capitolio de los EE.UU. bajo las aguas, como resultado de un posible caso de calentamiento global. Algunos profetizan movimientos ingentes de población al tornarse inhóspitas determinadas zonas del planeta. Otros sienten amenazada la soberanía y la seguridad nacional. El presidente Gorbachev ha declarado que, en el siglo que viene, la seguridad ecológica, y no la seguridad militar, constituirá el motivo principal de preocupación de todas las naciones, a medida que las condiciones medioambientales desencadenen catástrofes en todo el mundo.

Ahora bien, estos casos y supuestos deberían ir acompañados de una advertencia que rezara así: aunque los sucesos imaginados podrían quizás acontecer en determinadas circunstancias, no se trata de predicciones reales. ¿Qué aconseja, pues, la prudencia ante base tan poco firme? Parece claro que la actitud juiciosa sería la de tener en cuenta la debilidad de las razones esgrimidas sin que ello impidiese intervenir. Deberían tomarse primero las medidas avaladas por otros motivos de índole económica y ambiental, se produzca o no cambio climático [véase "Un clima cambiante" por Stephen H. Schneider; INVESTIGACIÓN Y CIENCIA, noviembre de 1989].

Después, a medida que el conocimiento científico fuera despejando las zonas de obscuridad, podrían emprenderse acciones más costosas si estuvieran justificadas; las decisiones políticas se hallarían, así, más estrechamente vinculadas al estado de los conocimientos. Se trata de una "política sin remordimientos", como algunos la llaman, o, en expresión lúdica, de "repartir las apuestas".

Un informe reciente del Consejo de Asesores Económicos consolida este enfoque. Establece que el coste de controlar las emisiones de dióxido de carbono y de tomar otras medidas destinadas a abordar el cambio climático se iría a los cientos de miles de millones de dólares. Puesto que tal reasignación de recursos evoca el espectro de graves consecuencias económicas, hemos de estar razonablemente seguros de que esas medidas valen lo que cuestan. Cualquier política nacional, "sin remordimientos", debería fijarse, entre sus principales objetivos, el fomentar que los gobiernos realizasen la inversión adecuada en un doble campo: la vigilancia global y la modelación matemática para extender nuestros conocimientos.

¿Por dónde empezar? La conservación de la energía y un mejor aprovechamiento de ésta, junto con la progresiva eliminación de la producción de CFC, habrían de ser las prioridades que atendiera la intervención nacional e internacional. El logro de un mayor rendimiento energético se justifica, en términos económicos, por sí solo. Mayor rendimiento que redundaría en una menor contaminación urbana y precipitación ácida. La sustitución de carbón y petróleo por gas natural, dentro del grupo de los combustibles fósiles, rebajaría apreciablemente las emisiones de dióxido de carbono por unidad térmica. Disponemos también de las técnicas necesarias para obtener un rendimiento mayor en la generación de electricidad y ahorrar en consumo de gasolina por kilómetro.

La prudencia recomendaría acometer inversiones importantes en fuentes de energía no fósiles. Las circunstancias favorecen nuevas y sustanciosas inversiones en energía nuclear pasivamente segura y aceptable por el público. Es razonable estimular el desarrollo de ciertas formas de energía solar —por medios fotovoltaicos o a través de la biomasa, por ejemplo—. La reforestación y la preservación de los bosques constituyen políticas acertadas que rinden muchos beneficios ecológicos y climáticos. Y valdría la pena poner los medios para investigar en cultivos resistentes al estrés y a las plagas.

Pero no hay que engañar a la gente. Esas medidas iniciales de la política sin remordimientos no resolverán el problema del calentamiento global. Su eficacia sólo lo retrasará en parte modesta, y el futuro podría exigir medidas más radicales. Con independencia de las que se tomen, no parece que haya posibilidad alguna de dete-

ner por completo el calentamiento global.

Las dificultades para alcanzar un acuerdo internacional sobre los procedimientos para mitigar el calentamiento global serán onerosas y duraderas. Tenemos una buena lección en las negociaciones para el Tratado de la Ley del Mar. Consumieron quince años y dieron por resultado un acuerdo que los EE.UU. no han firmado todavía, a causa de las exigencias, por parte de las naciones en vías de desarrollo, de recibir una sustancial ayuda tecnológica y económica. Cuestiones que adquirirán mayor complejidad y fuerza en cualquier "Tratado de la Ley Atmosférica".

Es probable que la humanidad tenga que adaptarse a algunos cambios climáticos. Y no se han estudiado bastante los modos de adaptación de la sociedad. Individuos, empresas y comunidades pueden acomodarse a los cambios climáticos de múltiples maneras. Los agricultores pueden optar por otros cultivos; podemos regular el uso del agua y modificar las prácticas de gestión.

Si se produjeran cambios climáticos de cierta entidad, necesitaríamos otras vías de adaptación. La subida del nivel del mar, que es una de las consecuencias predichas del cambio climático, podría inundar las regiones costeras bajas y provocar la filtración del agua salada, contaminando los cursos freáticos. Ante cuya eventualidad, la sociedad debería decidir entre invertir en estructuras de contención costera o adaptarse con nuevas estrategias de aprovechamiento del suelo. Los diques del Mar del Norte en Holanda son un ejemplo sobresaliente de adaptación a la elevación del nivel del mar.

Algunas adaptaciones requerirían un tiempo largo. Si el precio de la energía se disparase, habría que diseñar viviendas que en ese punto resultaran ahorradoras. Las actuales ciudades, con su gran dispersión suburbana, no lo son, y quizá debiéramos retornar a ciudades más compactas. En el supuesto de que decidiéramos seguir con la agricultura en las regiones secas, la sociedad tendría que pensar si valdrá la pena invertir en los sistemas de irrigación necesarios. A modo de ejemplo, el crecimiento económico de toda la parte occidental de los EE.UU. se ha basado en fuertes inversiones para el almacenamiento y transporte de agua de riego y de uso industrial.

Afortunadamente, por una vez, el tiempo podría estar de nuestro lado. Los gobiernos suelen actuar tan sólo cuando la amenaza se ha convertido en realidad. Intervienen ante amenazas militares o en regiones asoladas por desastres naturales. Si el clima cambia, lo probable es que lo haga gradualmente. Obligación nuestra sería descubrir precozmente el anegamiento de las costas por la frecuencia creciente de mareas altas y recorte de las franjas costeras. El propio calentamiento climático debería dejarse notar en una frecuencia, también creciente, de las olas de calor y otras anomalías meteorológicas. Lo probable es que los efectos de un calentamiento global tarden en 30 a 50 años en adquirir proporciones importantes, un lapso suficiente para poder tomar las medidas oportunas para hacer frente a la nueva situación.

¿Qué decir del debate que se desarrolla en las comunidades científica, ecologista y política? Nuestro entorno global está sometido a agresión desde muchos frentes. El calentamiento global no es más que una, quizá la más compleja, de esas cuestiones. Si es probable que los cambios que se producen en nuestra atmósfera acarreen consecuencias, debemos acotar los problemas y fomentar políticas sensatas para remediarlas. Insensato sería caer en excesos apocalípticos o negarlas a la manera del avestruz. Nos complacemos en creernos mucho más cultivados e ilustrados que las generaciones precedentes. En tanto no podamos enfocar los desafíos de nuestro ambiente de manera serena y objetiva, sin provocar la histeria pública ni excitar reacciones miopes e interesadas, no podremos pretender que lo seamos.

BIBLIOGRAFIA COMPLEMENTARIA

THERMAL EQUILIBRIUM OF THE ATMOSPHERE WITH A GIVEN DISTRIBUTION OF RELATIVE HUMIDITY. Syukuro Manabe y Richard T. Wetherald en *Journal of Atmospheric Science*, vol. 24, págs. 241-259; mayo de 1967.

CHANGING CLIMATE: REPORT OF THE CARBON DIOXIDE ASSESSMENT COMMITTEE. National Research Council, Carbon Dioxide Assessment Committee. National Academy Press, 1983.

THE GREENHOUSE EFFECT, CLIMATIC CHANGE, AND ECOSYSTEMS. Dirigido por Bert Bolin y Bo R. Doos. John Wiley & Sons, Inc., 1986.

GLOBAL TRENDS OF MEASURED SURFACE AIR TEMPERATURE. James Hansen y Sergej Lebedeff en *Journal of Geophysical Research*, vol. 92, D11, págs. 13345-13372; 20 de noviembre de 1987.

THE GREENHOUSE EFFECT: SCIENCE AND POLICY. Stephen H. Schneider en *Science*, vol. 243, n.° 4892, págs. 771-781; 10 de febrero de 1989.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Genes con homeobox y el plan corporal de los vertebrados

Esta familia de genes emparentados determina la forma del cuerpo. Subdivide al embrión a lo largo del eje cabeza-cola en campos celulares que se transformarán en miembros y otras estructuras

Eddy M. De Robertis, Guillermo Oliver y Christopher V. E. Wright

A partir de un óvulo fecundado de apariencia homogénea, surge gradualmente, por división celular, un embrión formado de piel, músculos, nervios y otros tejidos. Sin embargo, mucho antes de que la mayoría de las células del cuerpo naciente comiencen a especializarse, se halla establecido un esquema que define las principales regiones del cuerpo: la cabeza, el tronco, la cola, etcétera. Este esquema permite que combinaciones aparentemente idénticas de tejidos se dispongan en estructuras anatómicas claramente diferenciadas, tales como brazos y piernas.

Recientemente, los embriólogos han realizado grandes avances en el conocimiento de los mecanismos que controlan este proceso, otrora misterioso. En los últimos diez años, las poderosas técnicas de la biología molecular han permitido aislar y caracterizar genes particulares que intervienen en algunas de las decisiones relativas al establecimiento del esquema corporal embrionario. La clave reside en una familia de genes, conocidos como genes con homeobox (caja homeótica), que subdividen al embrión temprano en campos celulares con potencial para transformarse en tejidos y órganos específicos.

EDDY M. DE ROBERTIS, GUILLERMO OLIVER y CHRISTOPHER V.E. WRIGHT han trabajado en colaboración durante los últimos cinco años en la Universidad de California en Los Angeles (UCLA). De Robertis se doctoró en bioquímica por la Universidad de Buenos Aires en 1975. De 1980 a 1985 fue profesor de biología celular en el Biocentro de la Universidad de Basilea y, desde 1985, ocupa la cátedra Norman Sprague de química biológica de la facultad de medicina de UCLA. Oliver obtuvo su licenciatura en ciencias biológicas en la Universidad de la República Oriental del Uruguay en 1980 y su grado en biología molecular en la Universidad de México en 1984. Actualmente es profesor agregado de bioquímica en su *alma mater*, Uruguay. Wright, que se doctoró en bioquímica por la Universidad de Oxford, en 1984, enseña biología celular en la facultad de medicina de Vanderbilt.

El proceso de desarrollo de *Xenopus laevis*, una rana sudafricana, constituye un ejemplo adecuado sobre el modo en que un vertebrado toma forma. Los embriólogos modernos prefieren trabajar con este anfibio porque en toda época del año puede inducirse a la hembra a poner alrededor de 1500 huevos grandes y fácilmente fecundables. Dado que todos los vertebrados se desarrollan en forma similar, la mayoría de los mecanismos de la embriogénesis temprana de la rana son también aplicables en peces, pollos, seres humanos y otros animales.

Uno de los rasgos sobresalientes de las primeras fases del desarrollo es su rapidez. Una célula huevo fecundada de *Xenopus* se divide en dos al cabo de aproximadamente 90 minutos. A continuación, las células se dividen sincronizadamente cada 30 minutos, hasta alcanzar el número de 4000. En este estadio, el embrión se denomina blástula media y tiene la forma de una esfera hueca. A simple vista, todas las células parecen idénticas; no obstante, algunas de ellas, situadas alrededor del ecuador de la blástula media, ya están destinadas a volverse la capa celular llamada mesodermo. La formación del mesodermo viene inducida por factores de crecimiento proteicos liberados por las grandes células vitelinas situadas en el polo inferior del embrión.

La totalidad de la capa mesodérmica termina por desplazarse hacia el interior del embrión, durante el proceso de gastrulación. Hacia el fin de este proceso, se hallan delimitadas tres capas con potenciales de desarrollo diferentes: el mesodermo, el endodermo y el ectodermo. El mesodermo da origen a la mayor parte del cuerpo, incluyendo la columna vertebral, los músculos, los huesos y la pared del cuerpo. El endodermo genera la capa epitelial de tejido que reviste el tracto digestivo y otros varios órganos, tales como el pulmón, el hígado y el páncreas. El ectodermo da lugar a la piel y el sistema nervioso.

El ectodermo se transforma en sistema nervioso en respuesta a señales químicas difundidas a partir del mesodermo subyacente. Las señales inducen un engrosamiento de parte del ectodermo, que forma una estructura denominada placa neural. (En este estadio, el embrión se denomina néurula.) A continuación, los bordes de la placa neural se pliegan uno hacia otro, en tanto que el medio se hunde en el cuerpo embrionario. Finalmente, los bordes se fusionan para formar un tubo neural, que se convierte en la base del cerebro y la médula espinal.

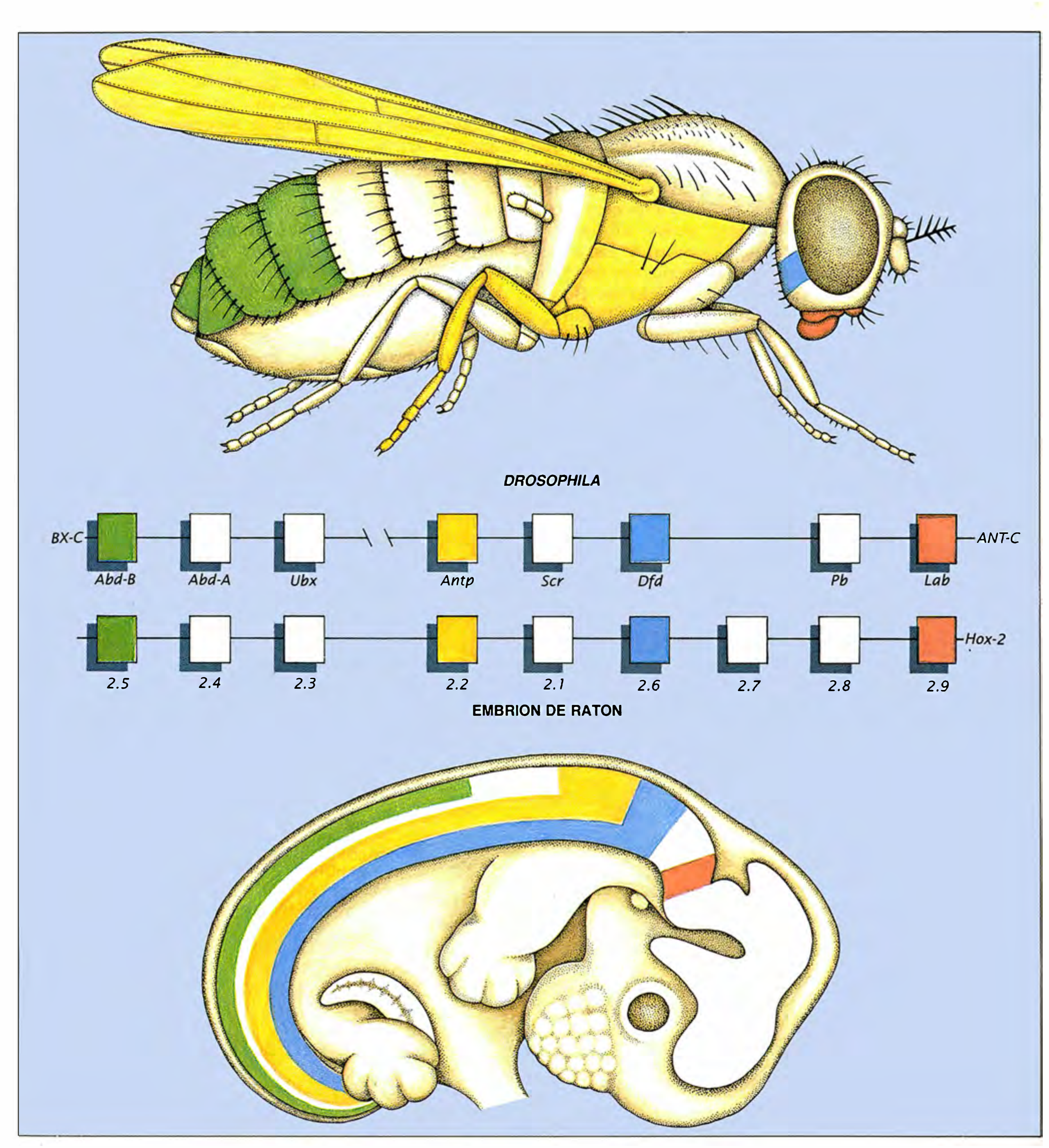
La determinación del eje anteroposterior (cabeza-cola) del embrión constituye la piedra angular del desarrollo, porque proporciona la línea central a lo largo de la cual se desarrollará el resto de las estructuras. Ross G. Harrison, de la Universidad de Yale, demostró que las células embrionarias están destinadas a transformarse en miembros y otras estructuras anatómicas específicas, muy poco después de completada la gastrulación.

En 1918, Harrison extrajo pequeños fragmentos de mesodermo de los flancos de néurulas de anfibios y los trasplantó en los flancos de otras. Cuando el tejido trasplantado provenía de una región determinada del donante, siempre daba origen a un

miembro anterior adicional en el receptor. Harrison observó que, si bien el mesodermo tenía en esos embriones la apariencia de una capa uniforme, las células de algún modo ya sabían a qué parte del cuerpo pertenecían.

Una de las peculiaridades notadas por Harrison fue que el embrión podía formar un miembro anterior, aun cuando él retirara todo el mesodermo que normalmente hubiera dado origen a dicho miembro. Interpretó este hallazgo en el sentido de que la zona de mesodermo circundante tenía también el potencial de inducir y controlar el crecimiento de los miembros. Esta extensa zona, con potencial para expresar una estructura, tomó el nombre de campo morfogenético.

Tras Harrison, numerosos científicos llevaron a cabo experimentos de trasplante en embriones de anfibios. Estos estudios permitieron establecer que el mesodermo constituye la capa celular crucial que especifica qué extremo del embrión es cabeza y cuál cola. El mesodermo de la néurula de anfibios se cartografió, o subdividió, en campos morfogenéticos para va-



1. LOS GENES CON HOMEOBOX controlan el desarrollo en animales tan dispares como *Drosophila melanogaster* (la mosca de la fruta) y el ratón. Estos genes dividen al embrión, a lo largo del eje cabeza-cola, en bandas con diferentes potenciales de desarrollo. La ubicación de un gen con homeobox en un cromosoma corresponde al lugar donde se expresa en el cuerpo: procediendo de izquierda a derecha, los genes controlan regiones del cuerpo más cercanas al extremo anterior del animal. Todos los genes con homeobox parecen tener un origen evolutivo común. En el presente diagrama se han coloreado en forma similar los genes con homeobox relacionados en *Drosophila* y en el ratón y las regiones corporales que controlan en cada animal. El mecanismo que determina la cabeza, el tronco y la cola puede haber surgido una sola vez en el curso de la evolución.

rios órganos: branquias, balancines (estructuras que aseguran que los renacuajos permanezcan erguidos), oídos, miembros anteriores, miembros posteriores, cola, etcétera. Dentro de cada campo morfogenético, el potencial de formación de órganos variaba gradualmente. En consecuencia, se postuló que cada campo morfogenético contenía un "campo de gradientes" de información, para determinar cada órgano. Como se explicará más adelante, el comportamiento de estos campos de gradientes se corresponde estrechamente con los modelos de expresión de determinados grupos de genes.

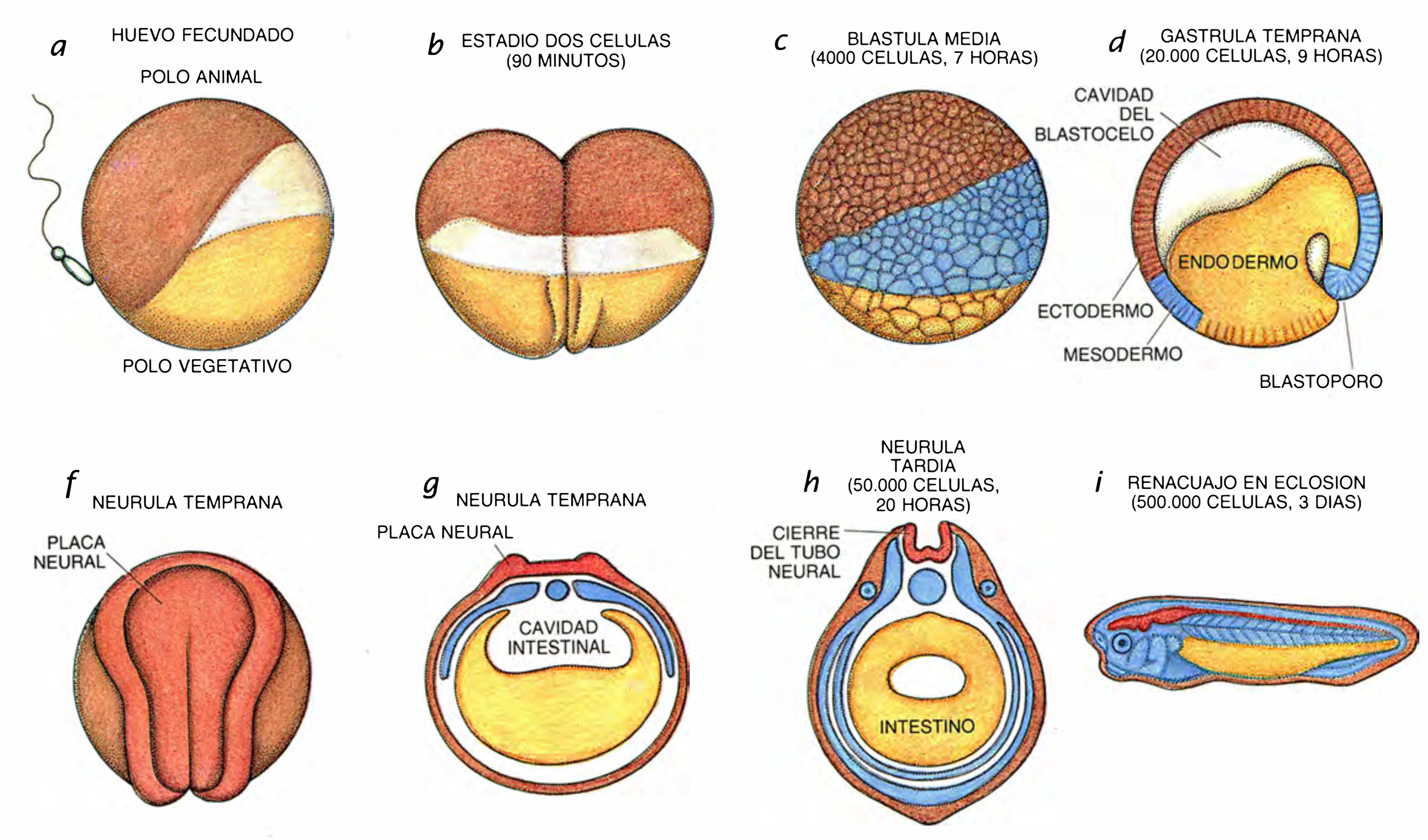
Los estudios mediante trasplantes, acerca del control ejercido por el mesodermo sobre el plan corporal de anfibios, terminaron en las postrimerías de la segunda guerra mundial, siendo reemplazados por estudios genéticos referidos a la manera en que el cuerpo adquiere forma. En 1948, Edward B. Lewis, del Instituto de Tecnología de California, inició un análisis genético profundo (que él continúa hasta hoy) de mutaciones homeóticas en la mosca de la fruta *Drosophila melanogaster*. La mutación homeótica provoca la sustitución de una parte del cuerpo por una estructura cuya ubicación normal correspondería a otro sitio. Por ejemplo, las moscas mutantes *bithorax* tienen dos pares de alas en vez de uno; las mutantes *Antennapedia* portan patas adicionales donde deberían poseer antenas.

Lewis descubrió que las transformaciones homeóticas podían deberse a mutaciones en genes individuales, aun cuando se hubieran necesitado cientos de genes activos para crear las alas y patas con la ubicación anormal. Por consiguiente, era razonable suponer que las mutaciones afectaban a genes rectores que dirigían la actividad de varios genes subordinados.

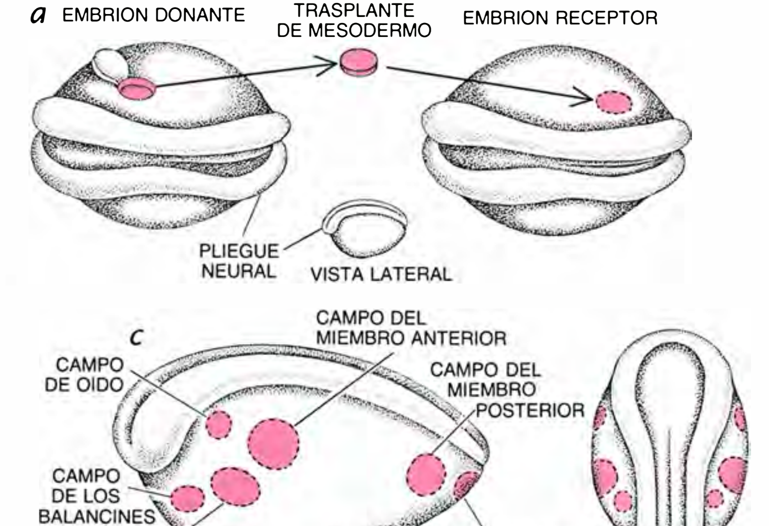
Una vez que la ingeniería genética dio a los científicos medios para aislar genes, se desató la carrera por descubrir y aislar genes homeóticos. A principios de 1980, David S. Hogness y Welcome Bender, de la Universidad de Stanford, aislaron los genes *Ultrabithorax, Abdominal-A* y *Abdominal-B* en el complejo *bithorax*. Walter J. Gehring y Richard L. Garber, del Biocentro de la Universidad de Basilea, y Matthew P. Scott y Thomas C. Kaufman, de la Universidad de Indiana, aislaron los genes del complejo *Antennapedia*, incluyendo los denominados *Labial, Proboscapedia, Deformado* y *Antennapedia.*

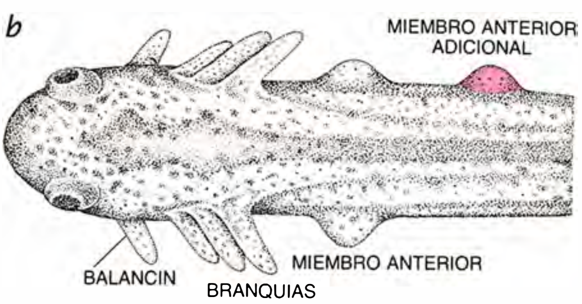
En 1983, se obtuvo un hallazgo crucial, cuando Gehring y su colega William J. McGinnis descubrieron que el gen *Antennapedia* contenía una secuencia de ADN que se encontraba también en otro gen rector del desarrollo. (A las secuencias similares de ADN en genes diferentes se les llama secuencias conservadas.) Dado que las secuencias conservadas de ADN pueden hibridarse, o aparearse, unas a otras, era posible marcar radiactivamente la secuencia conservada de ADN de *Antennapedia* y emplearla como sonda para localizar otros genes que contuvieran la misma región. De este modo, Gehring y McGinnis aislaron *Ultrabithorax, Deformado* y otros genes homeóticos. La región conservada de ADN fue identificada, en forma independiente, por Scott, quien se encontraba entonces en la Universidad de Colorado en Boulder.

Una nueva comprobación significativa obtenida por McGinnis fue que otros invertebrados, tales como los ciempiés y los gusanos de tierra, de los cuales se supone que han evolu-



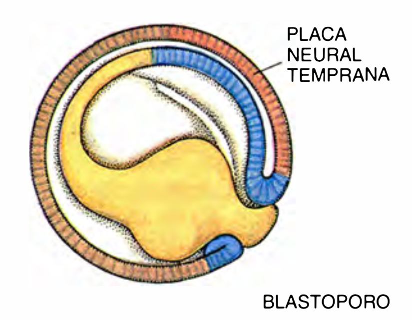
2. EL PLAN CORPORAL DE LOS VERTEBRADOS se establece mediante la formación inducida químicamente y el movimiento de capas celulares, como se ve en el desarrollo de *Xenopus laevis*, una rana sudafricana. A través de divisiones celulares rápidas (*a-c*), el huevo fecundado se transforma en una esfera de células hueca. Las grandes células vitelinas del polo inferior del embrión liberan factores de crecimiento que inducen la transformación de las células suprayacentes en la capa mesodérmica (*color azul*). El mesodermo constituye la capa crítica, que determina la polaridad anteroposterior del embrión. Otras dos capas celulares —el ectodermo (*color marrón*) y el endodermo (*color amarillo*)—, quedan establecidas durante la gastrulación, proceso mediante el cual el mesodermo emigra hacia el interior del embrión (*d-e*). En el estadio de néurula (*f-g*), el

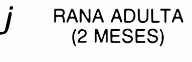


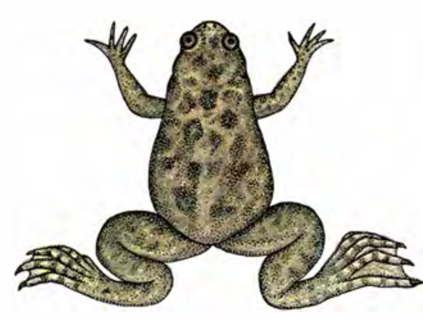


3. TRASPLANTES en que se injerta mesodermo de un embrión en estadio de néurula en otro embrión: tales trasplantes darán lugar a la formación de miembros u órganos adicionales. Si se trasplanta mesodermo de la región de los miembros anteriores, por ejemplo (*a*), el embrión receptor tendrá un miembro anterior supernumerario (*b*). Mediante tales experimentos en anfibios, los embriólogos han identificado campos morfogenéticos que determinan el desarrollo de diversas estructuras (*c*).

cionado los insectos, poseen también la misma región conservada de ADN. Resultaba clara la existencia de una relación entre las estructuras moleculares de numerosos genes cuyo control sobre el desarrollo celular embrionario se hallaba comprobado. La región conservada de ADN de







mesodermo induce la transformación de parte del ectodermo en placa neural (*rojo*). Como se ve en el corte transversal (*h-i*), la placa neural se cierra sobre sí misma y se convierte en tubo neural, predecesor del cerebro y de la médula espinal en el animal adulto (*j*). (P. J. Wynne.)

cada gen homeótico fue apodada "homeobox".

El homeobox codifica una secuencia de 60 aminoácidos, muy similar en los productos proteicos de la mayoría de los genes homeóticos. En la proteína, esa secuencia se conoce como el homeodominio. Su función consiste en reconocer y unirse a secuencias específicas de ADN en los genes regulados por genes homeóticos [véase "Base molecular del desarrollo", por Walter J. Gehring; INVESTIGACIÓN Y CIENCIA, diciembre de 1985.]

La cadena polipeptídica en el homeodominio consta de cuatro hélices, una de las cuales es responsable del reconocimiento de una secuencia específica de ADN. Debido a que esta hélice es casi idéntica en todas las proteínas con homeodominios, todas estas proteínas se unen a secuencias bastante similares de ADN. Al unirse a genes en una célula, las proteínas con homeodominios activan o reprimen la expresión de esos genes subordinados.

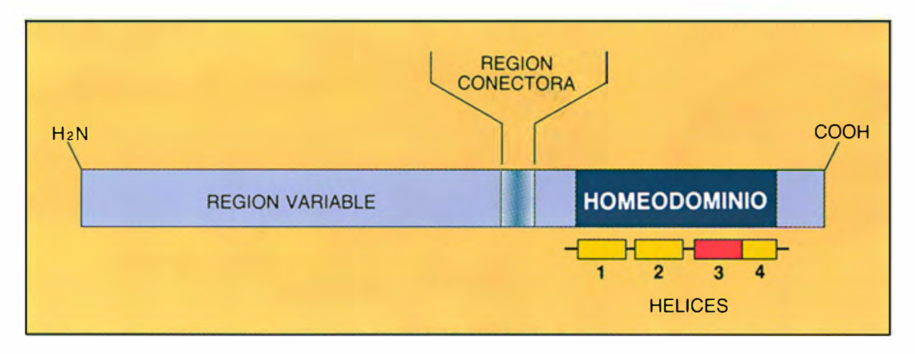
Nuestro trabajo de investigación sobre homeoboxes se inició en 1983, cuando uno de nosotros (De Robertis) tenía su laboratorio del Biocentro en el mismo piso que Gehring. Nos veníamos interesando en el desarrollo de *Xenopus laevis* desde hacía ya cierto tiempo [véase "Trasplante de genes y análisis del desarrollo", por Eddy M. De Robertis y J.B. Gurdon; INVESTIGACIÓN Y CIENCIA, febrero de 1980]. Al seguir los grandes avances realizados en los estudios sobre *Drosophila*, se nos hizo evidente que, si queríamos llegar a un conocimiento comparable del desarrollo de los vertebrados, deberíamos identificar en ellos los genes rectores.

Sin embargo, la falta de conocimiento acerca de la genética de las ranas parecía una barrera infranqueable para lograr nuevos avances. Aun cuando se disponía de estudios razonables sobre la genética de ratones, no existían candidatos verdaderos en lo concerniente a genes rectores de la embriogénesis.

Decidimos intentar lo que en aquel momento parecía un experimento loco: aislar un gen similar a *Antennapedia* del ADN de rana, con las sondas de homeobox de las moscas de la fruta de McGinnis y Gehring. Había pocas razones para creer que el ADN de rana contuviera ese gen o que los genes de especies tan poco emparentadas pudieran ser significativamente similares. No obstante, creímos que el intento valía la pena. Algunos de nuestros colegas eran escépticos en cuanto a que ese experimento pudiera funcionar y dos de nuestros alumnos declinaron su ayuda sobre esa base.

Pronto nos encontramos celebrando con una botella de champaña. En nuestro primer intento, trabajando con Andrés E. Carrasco, alumno de postdoctorado de nuestro laboratorio, tuvimos éxito. Analizamos la secuencia de ADN del gen de rana aislado en nuestro experimento, llamado actualmente *XlHbox 1*, y confirmamos que contenía la región del homeobox. Este descubrimiento sugería con fuerza que, por fin, podía disponerse de un gen que controlaba directamente el desarrollo de los vertebrados.

Mal podíamos imaginar, luego de nuestro primer experimento, que llevaría seis años más, y el esfuerzo de laboratorios de todo el mundo, confirmar que los genes con homeobox



4. PROTEÍNAS CON HOMEODOMINIOS, unidas al ADN, regulan la expresión génica. Se componen de una región variable, la que determina una actividad proteica específica, una pequeña región de articulación conectiva y un homeodominio, secuencia de sesenta aminoácidos, que es similar en todas las proteínas de este tipo. Esta secuencia es codificada por las regiones de homeobox de los genes. El homeodominio consta de cuatro hélices alfa (*1-4*), una de las cuales (*rojo*) reconoce y se une a una secuencia específica de ácido desoxirribonucleico en los genes blanco, o diana. (E. Bell.)

de vertebrados estaban directamente involucrados en el control del desarrollo. Aun así, los avances iniciales en el estudio de mamíferos fueron rápidos. Trabajando con ratones, Frank H. Ruddle, de la Universidad de Yale (entonces con licencia sabática en el Biocentro), y Peter Gruss, del Instituto Max Planck de Medicina Experimental de Gotinga, aislaron numerosos genes con homeobox. Dado Boncinelli, de la Universidad de Nápoles, obtuvo éxitos similares con genes humanos. Las proteínas codificadas por todos estos genes con homeobox se diferencian mucho unas de otras, excepto en la región altamente conservada del homeodominio.

La lista de proteínas con homeodominios creció en 1988, al lograrse la primera purificación de factores de transcripción. Estos factores son proteínas que aumentan la expresión de determinados genes diana. Al secuenciarse los factores de transcripción, se encontró que algunos de ellos contenían homeodominios, lo que indicaba que eran productos de genes con regiones de homeobox. Tales investigaciones bioquímicas confirmaron, en forma independiente, que los genes con homeobox regulaban la actividad de otros genes.

Pero, ¿de qué modo orquestan los genes con homeobox la diferenciación celular durante el desarrollo? La observación de las regiones donde las proteínas producto de los genes con homeobox se localizan en el cuerpo del embrión, en varios estadios del desarrollo, aporta un indicio a la respuesta. La proteína de *Xenopus XlHbox 1*, por ejemplo, se encuentra en una banda estrecha de células, situada detrás mismo de la cabeza del embrión de la rana. Esta banda está formada tanto por el mesodermo como por la médula espinal anterior y la cresta neural.

En estas capas tisulares, el límite anterior y el límite posterior de expresión de *XlHbox 1* guardan una clara alineación. Porque el mesodermo induce, según es sabido, las características anteroposteriores del tejido neural, parece posible que el mesodermo que expresa *XlHbox 1* induzca también la expresión del mismo gen en células de la placa neural suprayacente.

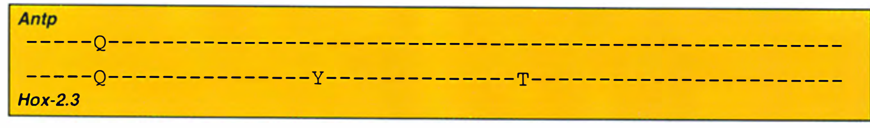
Otros genes con homeobox tienen actividad en regiones diferentes. Basándose en los patrones de expresión de los genes con homeobox, podemos concebir el embrión de vertebrados subdividido en campos celulares anteroposteriores con diferentes potenciales de desarrollo. Esta subdivisión del cuerpo embrionario precede a la formación de órganos o estructuras específicos.

Aun cuando los homeodominios codificados por diferentes genes con homeobox son muy similares entre sí, para identificarlos se puede recurrir a diferencias características en sus secuencias de aminoácidos. Algunos homeodominios se asemejan mucho más estrechamente que otros. Tomando como criterio estas similitudes y diferencias, resulta interesante destacar que determinados homeodominios de mamíferos se parezcan mucho a aquellos producidos por determinados genes de la mosca de la fruta.

Consenso
RKRGRTTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN

Labial
NNS---NF-NK-LT--------------A------NT-Q-N-T-V----------Q--RV
PGGL--NF-TR-LT---------K--S-A--V---AT-G-N-T-V----------Q--RE
Hox-2.9

Deformado
P--Q--A---H-I--------Y--------------T-V-S-----------------D-
P--S--A---Q-V--------Y---------V--------S-----------------DH
Hox-2.6



Antp
-----Q--
-----Q---------------Y--------------T-----------------------
Hox-2.3

Abd-B
VRKK-KP-SKF---------L--A-VSKQK-W-L-RN-Q-----V----------N--NS
SRKK-CP--K----------L--M----D--H-V-RL-N-S---V----------M--L-
Hox-2.5

5. TODOS LOS HOMEODOMINIOS son esencialmente similares, pero los producidos por genes de determinados insectos y mamíferos guardan un parecido muy estrecho. Aquí se muestran secuencias de aminoácidos de varios homeodominios. Los genes *Labial*, *Deformado*, *Antp* y *Abd-B* provienen de la mosca de la fruta, *Drosophila*; y los cuatro genes *Hox* análogos provienen de ratones. En cada secuencia, las letras representan aminoácidos. El guión indica que el aminoácido es el mismo que en la fila del consenso, un promedio de todas las secuencias de homeodominios. (E. Bell.)

Al analizar patrones de expresión de varios genes con homeobox de embriones de ratón, Robb Krumlauf, del Consejo de Investigaciones Médicas de Londres, y Denis Duboule, del Laboratorio Europeo de Biología Molecular de Heidelberg, realizaron importantes observaciones, cada uno por su lado. Con anterioridad, se había demostrado que, lo mismo en vertebrados que en invertebrados, los genes con homeobox se agrupaban en complejos, o grupos, en un cromosoma. En otras palabras, en la molécula lineal de ADN que compone un cromosoma, los genes con homeobox están dispuestos en un orden preciso, de izquierda a derecha.

Krumlauf y Duboule descubrieron inesperadamente que, en ratones, el orden de los genes con homeobox de un complejo se correspondía directamente con el lugar de expresión de los genes. Los genes con homeobox situados cerca del extremo izquierdo de un complejo se expresan en las regiones posteriores del cuerpo y los genes de la derecha se expresan más cerca

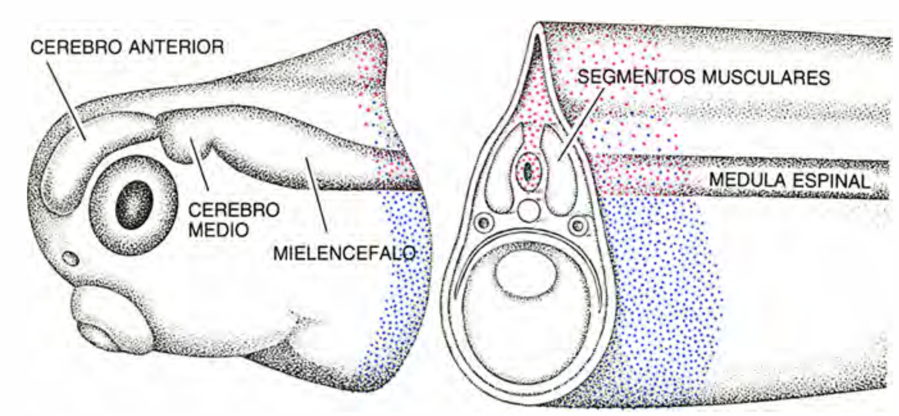
de la cabeza. Lewis había observado el mismo patrón en *Drosophila*, muchos años antes.

Todos los vertebrados poseen cuatro complejos de homeobox, cada uno de ellos localizado en cromosomas distintos. Es probable que estos complejos hayan surgido a lo largo de la evolución mediante la duplicación de un único grupo de genes con homeobox en invertebrados. Cada ser humano posee, pues, cuatro genes que se parecen, por ejemplo, al gen *Abdominal-B* de la mosca de la fruta y otros cuatro similares al *Deformado*.

Un principio unificador se aplica a todos los complejos de homeobox: los genes que se expresan en zonas posteriores se alojan a la izquierda y, a la derecha, los que se expresan en zonas anteriores. Así, en el ADN cromosómico, los genes con homeobox se disponen en el mismo orden en el que se expresan a lo largo del eje anteroposterior del cuerpo. Tal disposición extraordinaria puede haberse producido a consecuencia de que los genes con homeobox deban activarse en un orden determinado.

Se van acumulando las pruebas sobre cómo se produce este despliegue secuencial de genes con homeobox. En los embriones de vertebrados, el ácido retinoico (compuesto relacionado con la vitamina A, que en ocasiones puede provocar graves defectos en recién nacidos) y los factores de crecimiento peptídicos constituyen buenos candidatos para proveer esas pistas posicionales. Podrían transmitir esa información mediante la activación selectiva de genes con homeobox en el mesodermo, elemento clave en la determinación del plan corporal.

Mediante el agregado de ácido retinoico a células embrionarias en cultivo, el grupo de investigación de Boncinelli demostró que el compuesto activaba numerosos genes con homeobox. Douglas A. Melton, de la Universidad de Harvard, comprobó que el factor de crecimiento de fibroblastos (que induce la formación del mesodermo en el embrión temprano) activaba selectivamente los genes posteriores con homeobox en el embrión de rana. En nuestro laboratorio de la Universidad de California en Los Angeles, Ken W.Y. Cho mostró que una proteína parecida al TGFβ (factor de crecimiento transformante β) activaba sólo genes anteriores.



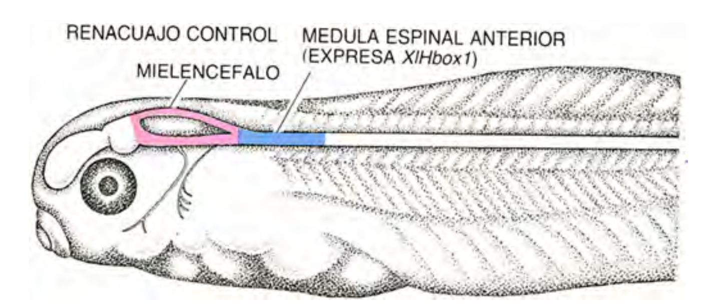
7. LOS GENES CON HOMEOBOX se expresan en bandas discretas a lo largo del eje anteroposterior del embrión. Por ejemplo, en el renacuajo de *Xenopus laevis*, el gen *XlHbox 1* se expresa en una región del tronco anterior del cuerpo. La proteína producida por el gen se encuentra en los núcleos celulares tanto de tejidos mesodérmicos (*azul*), como ectodérmicos (*rojo*). El miembro anterior crece enteramente a partir de células del mesodermo que expresan la proteína *XlHbox 1*.

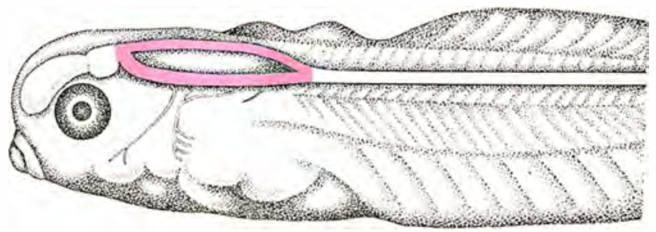
Una vez activados, ¿especifican los genes con homeobox las identidades y destinos de las células embrionarias, dando por consiguiente forma al cuerpo y dirigiendo la constitución de los órganos? ¿O son indirectos sus efectos? Los resultados de dos experimentos proporcionan argumentos en favor de un papel directo.

En el primer ensayo, inyectamos anticuerpos dirigidos contra la proteína *XlHbox 1* en embriones unicelulares de *Xenopus*. Los anticuerpos se fijaron a la proteína y la inactivaron durante el período crucial en el que se establece el plan corporal. Al examinar los renacuajos que se desarrollaron, descubrimos que los tejidos que normalmente expresaban *XlHbox 1*, y que deberían haberse transformado en una sección de la médula espinal anterior, se habían transformado, por contra, en estructuras del mielencéfalo. En efecto, la "pérdida de función" del *XlHbox 1* transformó parte de la médula espinal en una estructura más anterior.

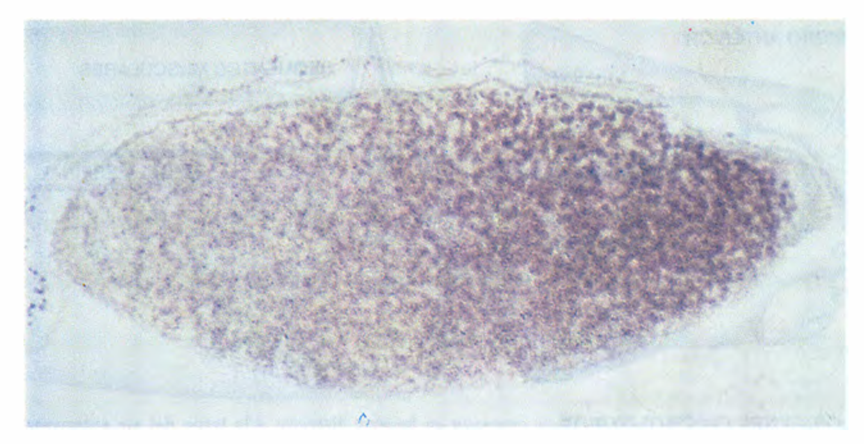
En el segundo ensayo, Gruss y Michael Kessel, del Instituto Max Planck, inyectaron ADN que contenía un gen con homeobox de ratón en embriones de ratón. El segmento de ADN se diseñó de manera que el gen con homeobox se expresara en la totalidad del cuerpo, incluso en regiones donde no lo hubiera hecho en condiciones normales, tales como la cabeza y el cuello. Los ratones resultantes presentaban frecuentemente defectos graves en la cabeza (verbigracia, fisura palatina). Más interesante aún, poseían una vértebra adicional y un disco intervertebral en la base del cráneo; algunos tenían un par de costillas supernumerarias en la región del cuello. Así, la "ganancia de función" de un gen con homeobox inducía transformaciones homeóticas exactamente similares a las observadas en las mutaciones de la mosca de la fruta.

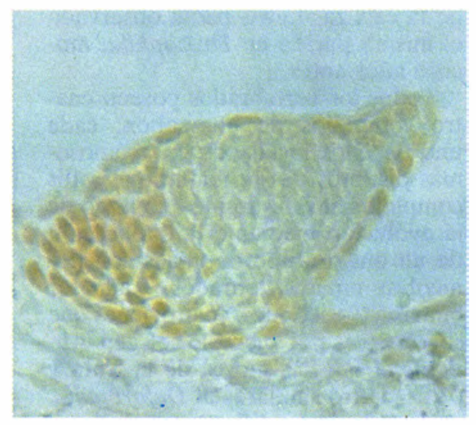
Otros trabajos indican también que los genes con homeobox cumplen un papel en la especificación de la identidad celular. Como antes se indicó, los genes con homeobox se expresan





6. INHIBICIÓN de una proteína con homeodominio. Dicha inhibición puede alterar el destino de los tejidos embrionarios. En renacuajos normales (*izquierda*), la proteína *XlHbox 1* se expresa en una región definida de la médula espinal anterior (*azul*). Si se inyectan anticuerpos contra la proteína *XlHbox 1* en un embrión unicelular, en el renacuajo resultante (*dibujo de la derecha*) aquella región se transformará en mielencéfalo (*rojo*).





8. EN ESTAS TINCIONES de brotes de miembros se pueden observar gradientes de proteínas con homeodominios. En el brote de ala de pollo (*izquierda*), la mayor concentración de proteína *Hox 5.2* se encuentra en los núcleos celulares cercanos a la región posterior (*borde derecho*). En el brote de aleta pectoral del pez cebra (*derecha*), la concentración más elevada de proteína *XlHbx 1* está en la región anterior (*borde izquierdo*). La aleta pectoral es el precursor del miembro anterior de los tetrápodos. Los gradientes son mecanismos eficaces para transmitir la información posicional.

intensamente en bandas a lo largo del eje anteroposterior, en la fase temprana del desarrollo. Más tarde, cuando los órganos de estas regiones se están formando, los mismos genes con homeobox se expresan de nuevo con intensidad. En estos estadios más tardíos, los genes con homeobox parecen proporcionar etiquetas moleculares que recuerdan a las células el lugar del embrión donde se originaron.

El desarrollo del miembro anterior constituye un caso particularmente informativo. La totalidad del campo del miembro anterior se deriva de la banda de mesodermo que expresa *XlHbox 1*. Las células proliferan en la banda y forman un brote (una insinuación) de miembro anterior que, en *Xenopus*, aparece hacia la tercera semana después de la fecundación.

En este estadio, el mesodermo del brote de miembro anterior parece uniforme, pero contiene un gradiente de la proteína de *XlHbox 1*. O sea, que la proteína abunda más en los núcleos celulares situados a lo largo del lado anterior del brote del miembro —el lado que da origen al dígito mayor— y menos en los núcleos del lado posterior, que da origen al dígito menor. Mientras el miembro crece y se conforma, la concentración de proteína *XlHbox 1* se mantiene en el nivel más elevado cerca del hombro, en el extremo proximal del brazo. Por contra, la proteína producto de otro gen, el *Hox 5.2*, establece un gradiente cuyo nivel máximo se encuentra a lo largo del lado posterior y el extremo distal del miembro, un patrón exactamente inverso al establecido para *XlHbox 1*.

En embriones de rana, pollo y ratón, se pueden detectar gradientes proteicos de *XlHbox 1* y de *Hox 5.2*. En el desarrollo de los miembros anteriores, también están comprometidos otros genes con homeobox. Duboule identificó otros tres genes con homeobox adyacentes al *Hox 5.2*, los que se activan secuencialmente a medida que la punta de la extremidad crece. El orden en que los genes se activan corresponde al orden que presentan en el ADN.

Gradientes de proteínas o de otras moléculas constituyen indicadores apropiados para especificar las posiciones celulares y para orientar con eficacia sus destinos. Por ejemplo, las células de una insinuación o brote de miembro pueden formar dígitos separados respondiendo, de manera diferente, a cantidades variables de una proteína. Sería menos económico el que una proteína diferente especificara cada dígito.

En conclusión, el análisis de los gradientes de proteínas con homeodominios durante el desarrollo de los miembros revela que la misma batería de genes con homeobox que establece el eje cabeza-cola vuelve a usarse más tarde para determinar las posiciones celulares durante el desarrollo de los miembros. Las proteínas con homeodominios se encuentran en los núcleos celulares, como se podría esperar de las proteínas combinadas de ADN que activan y desactivan genes. Aún se ignora la manera en que se establecen los gradientes de proteínas nucleicas en los brotes de extremidades. Probablemente estén implicadas señales de comunicación intercelulares, similares a las involucradas en la formación del eje y tal vez mediadas por factores de crecimiento o ácido retinoico.

Además de cuanto nos ha aclarado sobre el desarrollo embrionario, el estudio de los genes con homeobox ha aportado nuevas perspectivas sobre la evolución. Debido a que el orden de los genes con homeobox es similar en vertebrados e invertebrados, los primeros complejos con homeobox debieron haber evolucionado eones atrás, en platelmintos u otros organismos primitivos que fueron los ancestros comunes tanto de seres humanos como de insectos. Sería interesante saber si los organismos pluricelulares más primitivos con polaridad anteroposterior, tales como los rotíferos, poseen también complejos de genes con homeobox. La asombrosa conservación de los complejos a través de la evolución sugiere que, una vez encontrada una vía eficaz de determinación del eje anteroposterior, resultó más fácil producir nuevas formas corporales modificando ese sistema que creando estrategias totalmente nuevas.

La actividad de los genes con homeobox ofrece también claves para entender la posible evolución de estructuras anatómicas específicas. Durante mucho tiempo, los científicos se interrogaron, por ejemplo, sobre el origen del brazo, o del miembro anterior, en los tetrápodos. Debido a que los peces primitivos denominados celacantos poseían aletas pectorales con articulaciones óseas, durante muchos años se creyó que el brazo había evolucionado de las aletas pectorales.

Apoyo para esta teoría surge del trabajo de Anders Molven y Charles B. Kimmel, de la Universidad de Oregón. El *XlHbox 1* se expresa en primer lugar en una región circular del mesodermo lateral, correspondiente al campo de la aleta pectoral. En este estadio, la expresión del gen se corresponde exactamente con la de un campo morfogenético definido por Harrison en 1918.

A medida que las células proliferan, la proteína *XlHbox 1* forma un

agudo gradiente en el brote de aleta pectoral, similar al gradiente encontrado en insinuaciones de miembros anteriores de ranas, pollos y ratones. Este patrón sugiere que *XlHbox 1* es un gen antiguo, cuya función en el gradiente-campo del miembro antecede a la aparición de estructuras tetrápodas tales como los dígitos. Probablemente sea mucho lo que se pueda aprender abordando la embriología comparada de los vertebrados en el ámbito de la expresión génica.

A pesar de que llevaría largo tiempo entender con exactitud cómo cooperan entre sí los genes para organizar células que, a partir de un huevo aparentemente homogéneo, originen un renacuajo nadador, el estudio molecular del desarrollo de los vertebrados ya ha dado un gran salto adelante. La expresión de los genes con homeobox puede proporcionar una explicación molecular a los gradientes-campos reconocidos por los embriólogos experimentales hace muchas décadas. En el espectro zoológico, los genes que controlan el eje anteroposterior tienen un grado de conservación que sobrepasa las expectativas más delirantes de los investigadores. Se han identificado las moléculas que pueden estar implicadas en la transmisión de la información posicional, tales como el ácido retinoico y los factores de crecimiento. Ahora se plantea la posibilidad de investigar cómo cambia la forma del cuerpo en el transcurso de la evolución. Los estudiantes que comienzan hoy a trabajar en el laboratorio podrán un día ser capaces de responder preguntas simples del tenor siguiente: ¿en virtud de qué difiere un brazo de una pierna? ¡Qué buen momento para comenzar!

BIBLIOGRAFIA COMPLEMENTARIA

A GRADIENT OF HOMEODOMAIN PROTEIN IN DEVELOPING FORELIMBS OF *XENOPUS* AND MOUSE EMBRYOS. Guillermo Oliver et al. en *Cell*, vol. 55, n.º 6, págs. 1017-1024; 23 de diciembre de 1988.

THE MURINE AND *DROSOPHILA* HOMEOBOX GENE COMPLEXES HAVE COMMON FEATURES OF ORGANIZATION AND EXPRESSION. Anthony Graham et al. en *Cell*, vol. 57, n.º 3, págs. 367-378, 1989.

INTERFERENCE WITH FUNCTION OF A HOMEOBOX GENE IN *XENOPUS* EMBRYOS PRODUCES MALFORMATIONS OF THE ANTERIOR SPINAL CORD. Christopher V.E. Wright et al. en *Cell*, vol. 59, n.º 1, págs. 81-93; 6 de octubre de 1989.

VARIATIONS OF CERVICAL VERTEBRAE AFTER EXPRESSION OF A *Hox 1.1* TRANSGENE IN MICE. Michael Kesset et al, en *Cell*, vol. 61, n.º 2, págs. 301-308; 20 de abril de 1990.

El colisionador LEP

El Gran Colisionador Electrón-Positrón producirá un millón de partículas Z^0 este año. La abundancia de datos permitirá a los físicos someter a prueba, mejor que nunca, su conocimiento de las leyes de la naturaleza

Stephen Myers y Emilio Picasso

Aunque era media tarde del 14 de julio de 1989, festivo y, además, bicentenario de la revolución francesa, ingenieros e investigadores corrieron a la sala de control del Laboratorio Europeo de Física de Partículas (CERN) en Prevessin, Francia, en la frontera con Ginebra. Se había anunciado que el Gran Colisionador Electrón-Positrón (LEP), del laboratorio, estaba preparado para acoger el primer haz. Querían ser testigos del histórico suceso. El director general del CERN, Carlo Rubbia, y su predecesor Herwig F. Schopper, que había ofrecido todo su apoyo al LEP en su etapa de responsable, se encontraban entre los anhelantes espectadores.

A las cuatro menos cinco, el primer enjambre de positrones acabó su recorrido inaugural por el acelerador gigante en forma de anillo. Cada uno de los cientos de miles de componentes dispuestos a lo largo de los 26,7 kilómetros de circunferencia de la máquina funcionaba a la perfección.

Un mes después, en la misma sala de control, si bien menos concurrida, el personal trabajó hasta tarde. Se aprestaban para provocar la primera colisión de dos haces que circulaban en sentido opuesto; uno era de electrones; el otro, de positrones. Cinco minutos antes de la medianoche, se desconectaron las placas electrostáticas que mantienen los haces separados y se dejó que los haces chocaran en los centros de los cuatro detectores gigantes del LEP: Aleph, Delphi, L3 y Opal.

Diez interminables minutos después sonó el teléfono de la sala de control. Era el grupo del detector Opal, que llamaba para anunciar la detección, clara, de la primera partícula Z^0 producida por el LEP. Durante el tiempo que se tardó en llevar una fotocopia del suceso a la sala de control, el teléfono sonó para anunciar cinco posibles sucesos más. En medio de la euforia subsiguiente, Rubbia, que compartió el premio Nobel en 1984 por su papel en el descubrimiento del Z^0, recordó que él había ido a Estocolmo respaldado por sólo cinco sucesos.

En los cuatro meses subsiguientes, los detectores del LEP observaron más de 100.000 sucesos Z^0. Observaciones que han rendido ya un resultado importante: pruebas de que quizá haya sólo 12 partículas elementales, ordenadas en tres "generaciones". Se han publicado una treintena de artículos basados en los sucesos Z^0 registrados en 1989.

Se espera que el LEP genere este año en torno a un millón más de Z^0. Semejante aluvión de datos permitirá a los físicos investigar modos de desintegración insólitos, determinar con exactitud la masa y la vida media del Z^0, estudiar los procesos que rigen la producción de hadrones (partículas formadas por quarks) y comprobar sutiles aspectos de la teoría imperante. No menor importancia reviste el afán de los experimentadores del LEP por descubrir nuevos estados de la materia, sobre todo, el bosón de Higgs y quizá también el quark cima ("top"). Estas dos partículas dotadas de masa, aunque predichas por la teoría, no se han observado todavía. Los experimentadores buscarán también fenómenos que no están predichos por las teorías existentes.

El colisionador LEP está destinado a ser el instrumento más importante de la física de partículas del próximo decenio. La forma espectacular en que ha funcionado hasta la fecha es un tributo al ingenio y al esfuerzo decidido de los centenares de técnicos, ingenieros e investigadores de más de 25 países que han colaborado en su diseño y en el de sus detectores.

Visto desde el aire, el LEP es un gran octógono, casi circular, a caballo de la frontera franco-suiza. El anillo consta de ocho secciones rectas, de 500 metros de largo cada una, unidas por ocho arcos de 2,8 kilómetros de longitud. Los componentes del acelerador están en un túnel subterráneo de unos cuatro metros de ancho a una profundidad media de 100 metros.

Con un diámetro medio de 8486 metros y una energía de colisión de los haces de 100.000 millones de electronvolt, o gigaelectronvolt (GeV), el LEP es el mayor anillo de almacenamiento electrón-positrón y el más potente que se haya construido hasta ahora. Marca la culminación de una historia que empezó hace 30 años, en marzo de 1960, cuando Bruno Touschek dio un seminario en el Laboratorio Nacional Italiano de Frascati, en el que razonó la importancia de estudiar las colisiones entre los electrones y sus equivalentes de antimateria, los positrones. Propuso un acelerador en el que haces de electrones y posi-

STEPHEN MYERS y EMILIO PICASSO pertenecen al Laboratorio Europeo de Física de Partículas (CERN), instalado cerca de Ginebra. Myers se recibió de doctor en 1972 por la Universidad Queens de Belfast; ese mismo año se trasladó al CERN. Desde 1979 ha intervenido de manera destacada en aspectos importantes del proyecto, construcción y puesta en marcha del LEP. Actualmente es el subdirector de la división del Superprotón Sincrotrón-LEP. Picasso se doctoró en física por la Universidad de Génova en 1956 y pasó al CERN en 1964. Trabajó en comprobaciones precisas de la electrodinámica cuántica y de la relatividad especial. De 1981 a 1989 dirigió el proyecto LEP.

1. GRAN ANILLO DE ALMACENAMIENTO electrón-positrón, de 26,7 kilómetros de circunferencia. Se encuentra en un túnel a 100 metros bajo tierra. Paquetes de electrones y positrones que giran en sentido opuesto chocan a energías cercanas a 100.000 millones de electronvolt. Abajo se muestran, de izquierda a derecha, los imanes de focalización del haz, un sistema de transporte en el túnel y una hilera de cavidades de radiofrecuencia que impulsan los haces.

trones de la misma energía se dirigieran en sentidos opuestos siguiendo una trayectoria circular y chocaran frontalmente. Los productos de las colisiones se estudiarían a fin de obtener información sobre las leyes fundamentales de la naturaleza.

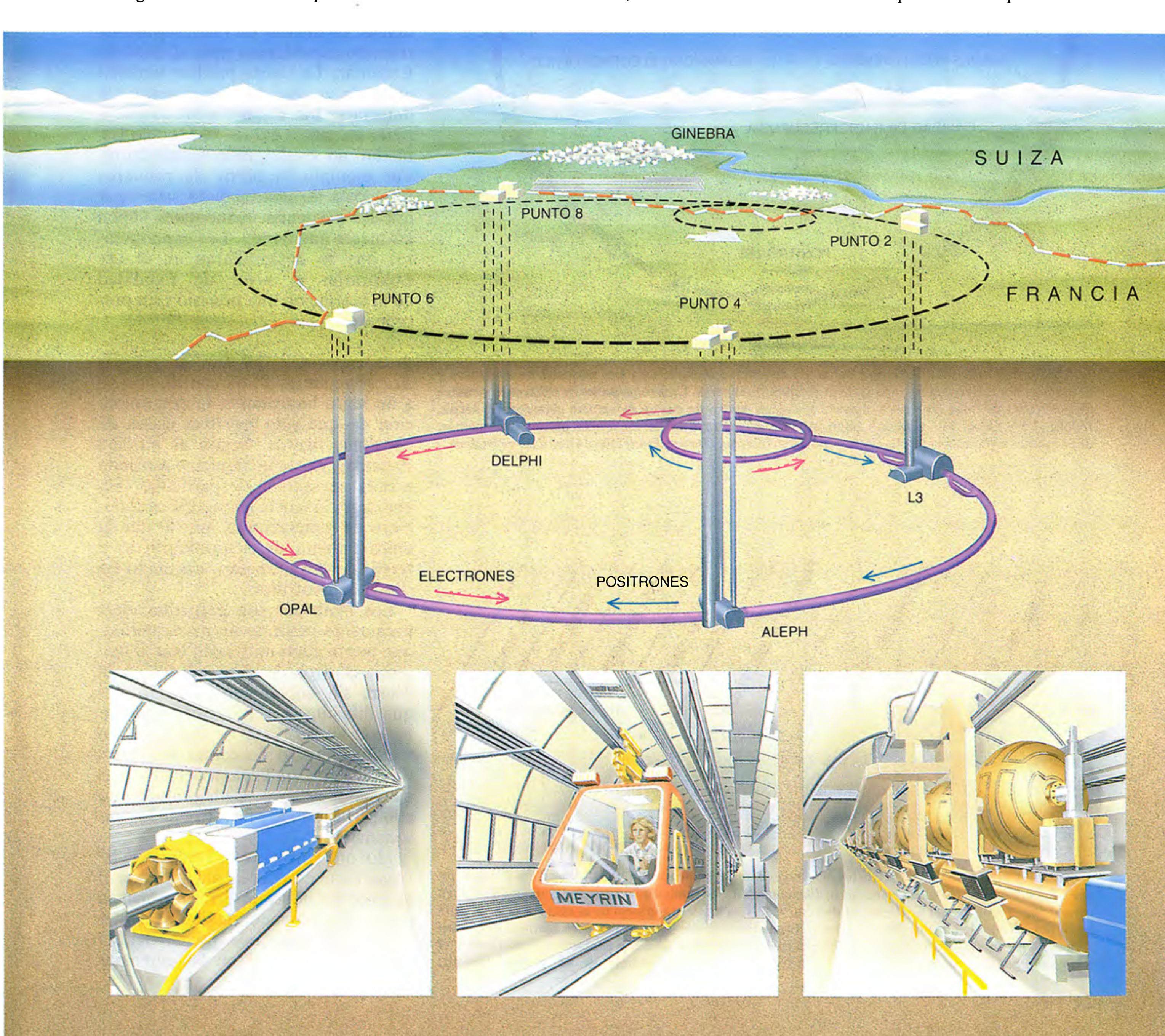
El primer anillo de almacenamiento de electrones y positrones se construyó en Frascati; se le llamó ADA ("anello di accumulazione"). El ADA medía 1,6 metros de diámetro y producía una energía máxima de haz de 250 millones de electronvolt (MeV). A esta máquina pionera le siguieron muchos anillos de almacenamiento de electrones y positrones, de mayor tamaño y con mayores energías, en la URSS, Francia, los EE.UU, Italia, Alemania Occidental y Japón.

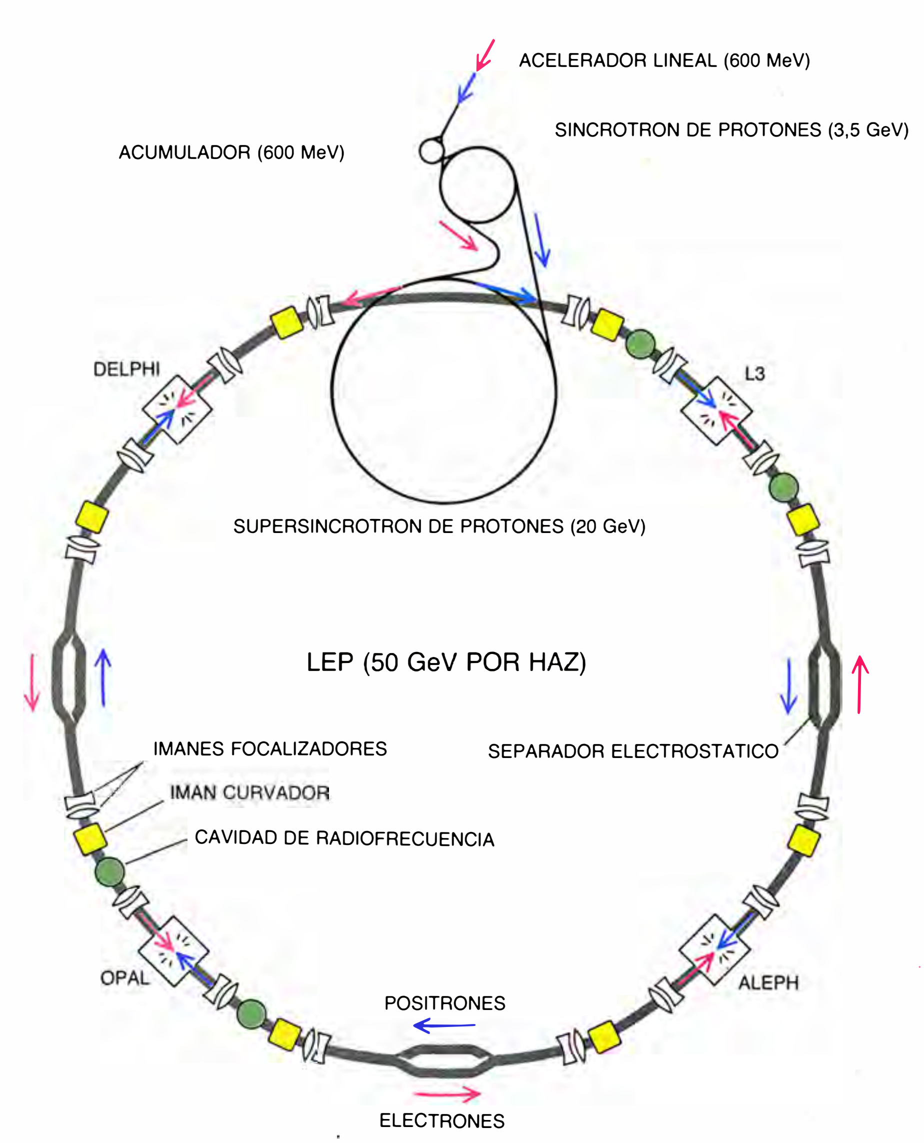
Todas estas máquinas se proponen crear nuevas partículas a partir de la energía liberada en la aniquilación del electrón por su antipartícula. De acuerdo con la fórmula fundamental de Einstein, $E = mc^2$, cuanto mayor sea la energía liberada, de mayor masa estará dotada la partícula creada. Como los electrones y los positrones tienen la misma energía y chocan frontalmente, la energía total disponible en el centro de masas del sistema es sencillamente el doble de la energía del haz. El LEP, con las actuales energías de colisión de hasta 110 GeV, genera partículas Z^0. Está previsto incrementarlas en el futuro a 200 GeV para que genere partículas W^+ y W^-. Cada una de estas tres partículas pesa unos 90 GeV, unas 100 veces la masa del protón.

Para entender el significado de estas partículas nos debemos referir a la teoría dominante de la materia y las fuerzas: el "Modelo Estándar". De acuerdo con el mismo, existen dos tipos básicos de partículas: fermiones y bosones. Los fermiones son los constituyentes básicos de la materia. Los bosones, "transmisores de la fuerza", se intercambian entre los fermiones generando atracción y repulsión entre las distintas clases de fermiones.

Hay dos clases de fermiones: los leptones y los quarks. Los leptones constan de partículas cargadas, como los electrones y los muones, y de neutrinos, partículas sin carga y apenas masa. Los quarks son los constituyentes de los hadrones; se encuadran entre éstos el protón y el neutrón.

Los fermiones se disponen en generaciones, cada una formada por dos quarks y dos leptones. La primera generación está formada por los quarks "arriba" y "abajo" (que forman los protones y neutrones), el electrón y el neutrino electrónico. La segunda generación comprende los quarks en-





2. VISTA A VUELO DE PAJARO del anillo de almacenamiento LEP. Los paquetes de electrones y positrones son acelerados por una serie de máquinas menores. Cuatro paquetes de electrones y cuatro de positrones circulan en sentido opuesto y chocan dentro de cuatro detectores gigantescos: Aleph, Delphi, L3 y Opal. Los paquetes también cruzan sus trayectorias en cuatro puntos intermedios, donde placas separadoras electrostáticas evitan su colisión. (Ilustración realizada por George Retseck.)

	MASA DEL Z^0 (GeV)	ANCHURA DEL Z^0 (GeV)	FAMILIAS DE NEUTRINOS	QUARK CIMA	BOSON DE HIGGS	SELECTRON	W WINO
				LIMITE INFERIOR DE LA MASA (GeV)*			
ALEPH	91.182 ± .056	2.541 ± .056	3.01 ± .20	45.8	24.0	44.2	46.0
DELPHI	91.171 ± .060	2.511 ± .065	2.97 ± .26	44.0	14.0	42.0	45.0
L3	91.160 ± .054	2.539 ± .054	3.29 ± .17	VEASE EPIGRAFE		41.0	44.0
OPAL	91.154 ± .051	2.536 ± .045	3.09 + .25 − .31	44.5	25.0	43.4	45.0

* NIVEL DE CONFIANZA DEL 95 POR CIENTO

3. RESULTADOS RECIENTES de los detectores del LEP. La masa y la "anchura" de la Z^0 implican que sólo hay tres neutrinos ligeros. No se dispone aún de los límites inferiores de las masas del quark cima ("top") y del bosón de Higgs del detector L3. (Los límites inferiores de la masa del quark cima están por debajo de los obtenidos por los colisionadores protón-antiprotón, pero también dependen menos de los modelos teóricos de su desintegración.) Los datos del LEP descartan el selectrón y el wino (compañeros supersimétricos del electrón y el *W*) por debajo de 44 gigaelectronvolt.

cantado y extraño, el muon y el neutrino muónico. Pertenecen a la tercera generación el quark "fondo" (bottom), el quark no observado "cima" (top), el tau y el neutrino tauónico.

Según el Modelo Estándar, los bosones "transmiten" las cuatro fuerzas fundamentales de la naturaleza. De transmitir el electromagnetismo se ocupan los fotones (partículas de luz). La fuerza fuerte, que liga los protones y neutrones en los núcleos atómicos, se transmite por los gluones. La gravedad es transmitida por los gravitones. El Modelo Estándar se apuntó un éxito al unificar el electromagnetismo y la fuerza débil, que es responsable de la desintegración radiactiva. En virtud de ello, la teoría predijo que la fuerza débil se transmitiría por los bosones vectoriales intermediarios: Z^0, W^+ y W^-.

El descubrimiento de estas tres partículas, en el CERN en 1983, supuso un rotundo espaldarazo para el Modelo Estándar. La teoría predice también el modo en que pueden generarse y detectarse los Z^0, W^+ y W^-. Mas, a pesar de tantos éxitos, el Modelo Estándar no es del todo satisfactorio. Por ejemplo, a partir de primeros principios no nos es dado saber por qué las partículas elementales tienen las masas que tienen. Los experimentos que se desarrollarán en el LEP explorarán el nivel de fiabilidad de la teoría con una precisión sin precedentes.

La capacidad del LEP de generar abundantes Z^0 ha contribuido ya a un logro importante: la determinación de que sólo hay tres clases de neutrinos ligeros. Según el Modelo Estándar, existe un neutrino asociado a cada generación de partículas. Sin embargo, la teoría no especifica el número de generaciones, por lo que la única manera de determinarlo es a través del experimento, contando las clases de neutrinos.

Los neutrinos son partículas eléctricamente neutras, sin masa apenas, que interactúan muy poco con la materia; chorros de ellos atraviesan la Tierra sin impedimento alguno. Lo que dificulta su observación directa. El colisionador LEP proporciona un medio indirecto de contar las clases de neutrinos, sin llegar a detectar esas partículas evasivas. Lo hace midiendo con precisión la "anchura" de la "resonancia" Z^0.

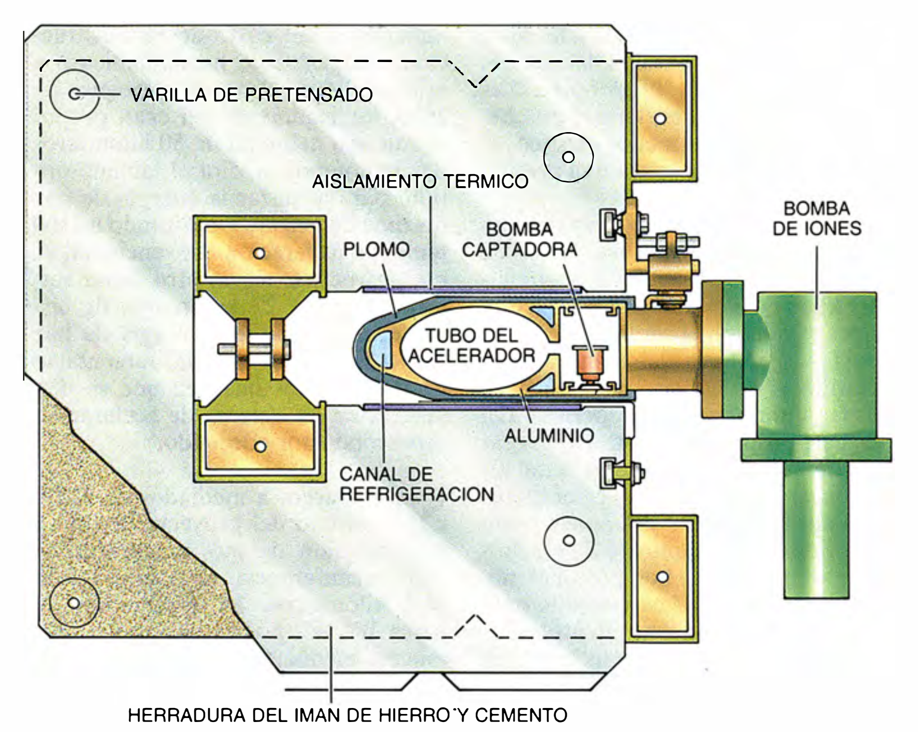
¿En qué consiste esta resonancia Z^0 y qué tiene que ver su anchura con el número de neutrinos? La resonancia alude al fenómeno siguiente: a medida que la energía de colisión del LEP crece gradualmente en las cercanías

de 91 GeV, aumenta el número de sucesos Z^0. Dicho número alcanza un pico, o "resuena", en torno a 91,2 GeV, para disminuir después cuando la energía sigue aumentando. El histograma del número de Z^0 en función de la energía del haz dibuja una curva acampanada.

La anchura de esta curva constituye la prueba crucial que limita el número de posibles partículas elementales. ¿Cómo? En realidad, la anchura constituye un parámetro de la incertidumbre en la energía, que, según el principio de incertidumbre, guarda una relación inversa con el tiempo: cuanto más ancha es la curva, más corta será la vida media de la partícula. A su vez, la vida media indica el número de modos de desintegración que están a disposición de la Z^0: cuanto mayor sea la cifra de maneras de desintegrarse la Z^0, más corta será su vida media y más ancha la curva. Por tanto, el número de clases de neutrinos afectará a la anchura.

Las mediciones de la anchura de la Z^0 por los cuatro detectores del LEP abonan la existencia de sólo tres clases de neutrinos ligeros. Queda definitivamente descartado un cuarto neutrino ligero. Las futuras teorías tendrán que tomar en consideración este hallazgo. (La resonancia Z^0 no revela si hay neutrinos pesados cuya masa supere la mitad de la de Z^0. Tales neutrinos serían miles de millones de veces más pesados que los neutrinos conocidos.) Conclusión ésta que deja sentir su efecto en cosmología y astrofísica. Por ejemplo, hemos de abandonar la posibilidad de que un cuarto neutrino que pese hasta 10 GeV aporte la "masa perdida" que se necesita para detener la expansión del universo.

Los investigadores del LEP están haciendo esfuerzos coordinados para colocar otras piezas del rompecabezas del Modelo Estándar en su sitio. Se están buscando dos partículas predichas por la teoría, el quark "cima" y el bosón de Higgs (que aparece en versión neutra y quizá también cargada). Los datos de la física de partículas, combinados con la precisa masa del Z^0, han llevado a los teóricos a predecir una masa del quark "cima" de unos 139 GeV, 30 GeV arriba o abajo. Hasta ahora, los experimentos del LEP han descartado un quark cima que pese menos de unos 45 GeV y un bosón de Higgs neutro que pese menos de 25 GeV. Si las partículas no se encuentran a las energías que el LEP es capaz de producir, ello puede querer decir simplemente que éstas pesan más, pero también puede indicar que el Modelo Estándar tiene fallos fatales.



4. TUBO DEL ACELERADOR, imán curvador y sistema de bomba de vacío, en un corte transversal. Los haces de partículas viajan por el centro del tubo. La pared de aluminio está revestida con plomo para evitar que la radiación de sincrotrón salga al exterior y genere gases corrosivos. Rodeando el tubo hay un gran imán curvador; como sólo tiene que generar un campo débil, se rebajó el coste del imán separando las placas de hierro con cemento. Para alcanzar un alto vacío de 10^{-11} torr, se emplearon bombas "captadoras" que reaccionasen con moléculas residuales y las retuviesen.

Al proyectar el LEP para que cumpliera sus objetivos científicos, los investigadores hubieron de considerar numerosos factores. El más crucial en la determinación del tamaño del anillo de almacenamiento LEP era el fenómeno de la radiación de sincrotrón. Toda partícula cargada que recorra una trayectoria circular radiará fotones (partículas de luz) y perderá una fracción de su energía. Si esta energía no se repone, las partículas se desacelerarán rápidamente y describirán una espiral hacia la pared del acelerador. A una energía de haz de 55 GeV, el LEP radia unos 200 millones de electronvolt por vuelta y necesita 16 megawatt de potencia a 350 megahertz. Esta potencia se suministra mediante una serie de cavidades de radiofrecuencia (RF). A su energía futura de 100 GeV, los haces del LEP radiarán más de 2800 megaelectronvolt por vuelta.

Así, pues, el coste que implica hacer funcionar un anillo de almacenamiento de electrones y positrones viene determinado por el coste de reponer la energía perdida a través de la radiación de sincrotrón. Para una partícula que recorra una trayectoria circular, esta pérdida es proporcional a la cuarta potencia de la energía de la partícula dividida por el radio de la circunferencia. Resulta claro que, para una energía de haz dada, se puede reducir la pérdida aumentando el radio de la circunferencia. Con igual claridad, resulta más caro construir un anillo mayor. En términos económicos, hay que lograr un equilibrio entre el coste de un anillo mayor y el coste de construir y hacer funcionar el sistema de RF.

La radiación de sincrotrón, aunque sea perniciosa para el coste de mantener la energía del haz, presenta también efectos positivos. La radiación "amortigua" las oscilaciones de las partículas individuales del haz. Cuando se inyectan nuevas partículas en el haz, sus energías oscilan en torno al valor central del haz. Cuanto mayores sean las oscilaciones, más deprisa las reducirá el amortiguamiento de la radiación. Por tanto, el amortiguamiento de la radiación permite acumular muchas partículas mediante la inyección repetida de nuevas partículas en el haz.

La fuerza de amortiguamiento, que, por sí sola, reduciría las oscilaciones de las partículas a cero, se contrarresta con una fuerza difusora que provoca que las partículas aumenten su amplitud de oscilación. Por tanto, el haz alcanza un estado de equilibrio en el que se compensan las dos fuerzas opuestas. La situación se asemeja a un globo. La presión del aire en el globo es análoga a la fuerza difusora,

y la elasticidad de la goma lo es al amortiguamiento de la radiación. Si la presión aumenta, el globo se dilatará, mientras que, si el balón está hecho de un material menos elástico, el tamaño será menor, para una presión determinada.

Dicho estado de equilibrio proporciona, y eso es admirable, haces de partículas casi insensibles a pequeñas perturbaciones, como son las fluctuaciones rápidas en el campo magnético director, a imagen del globo que se deforma y vuelve prestamente a su forma anterior. A las energías del LEP, esta propiedad sólo se da para los electrones y positrones, no para los protones. El físico de aceleradores halla ahí la principal diferencia entre anillos de almacenamiento de electrones y de protones. Los protones tienen una "memoria" casi infinita para cualquier perturbación, mientras que los electrones las "olvidan", por lo común en centésimas de segundo.

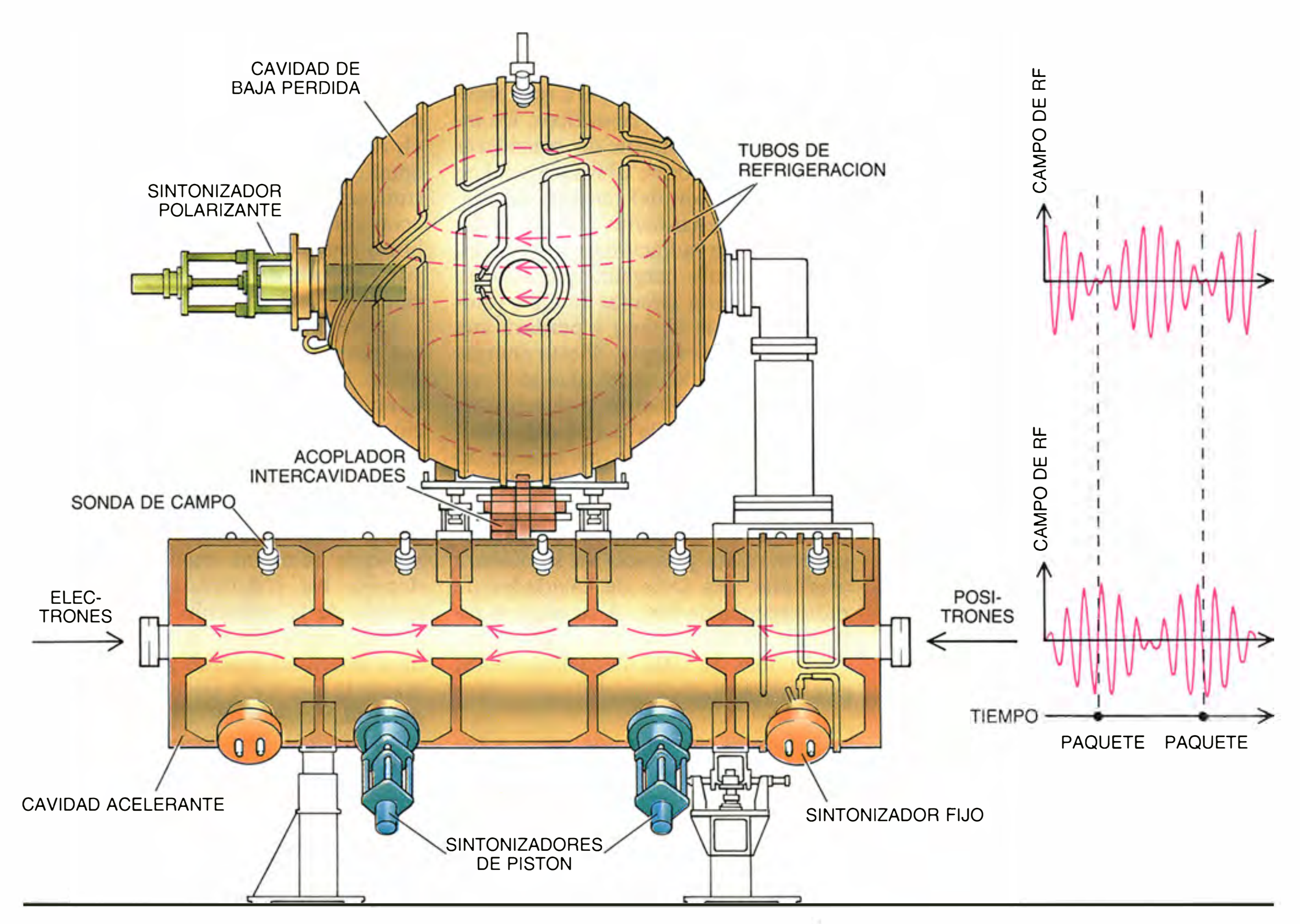
Los diseñadores del LEP tenían que producir una máquina que pudiera alcanzar la energía y la luminosidad (intensidad del haz) necesarias y equilibrar estas exigencias con las consideraciones del coste de la construcción, los gastos de funcionamiento y la flexibilidad futura. A principios de 1976, los ingenieros del CERN propusieron una máquina de 50 kilómetros de circunferencia. Era el tamaño óptimo para alcanzar la energía de haz máxima de 100 GeV utilizando un sistema de aceleración convencional. A ese proyecto le siguió otro de un anillo menor, de 22 kilómetros de circunferencia, con una energía de haz de 70 GeV, que podría aumentarse hasta 100 GeV una vez que se dispusiera de un sistema de aceleración superconductor adecuado.

Sin embargo, a mediados de 1979, el estudio del proyecto llegaba a la conclusión de que era preferible una circunferencia algo mayor, de 26,7 kilómetros. La producción de pares W^+/W^- podría alcanzarse, entonces, mediante sistemas de aceleración a temperatura ambiente y no tendría que depender del desarrollo de tecnología superconductora.

En diciembre de 1981 el consejo del CERN aprobó el proyecto. El colisionador funcionaría inicialmente a una energía máxima en centro de masa (energía total de colisión frontal) de 110 GeV (fase uno) y más tarde se desarrollaría para alcanzar energías de hasta 200 GeV (fase dos), con la posibilidad de aumentar aun más la energía hasta unos 240 GeV. La segunda fase permitirá estudiar la intensidad de la interacción entre las Z^0, W^+ y W^-, que es crucial para establecer la validez de la teoría electrodébil.

El trabajo de construcción empezó al poco de seleccionar el proyecto. Desde 1983 hasta 1988, el LEP constituyó el proyecto europeo de ingeniería civil más ambicioso. La ingeniería y la construcción de la infraestructura costaron, solas, más de la mitad del presupuesto total. El túnel del anillo principal constituye la parte más impresionante de la obra, a pesar de que representa sólo menos de la mitad de los 1,4 millones de metros cúbicos de roca y tierra que se han excavado. El resto de la obra subterránea incluye cuatro departamentos que albergan los detectores, 18 pozos y unas 60 cámaras y cavidades.

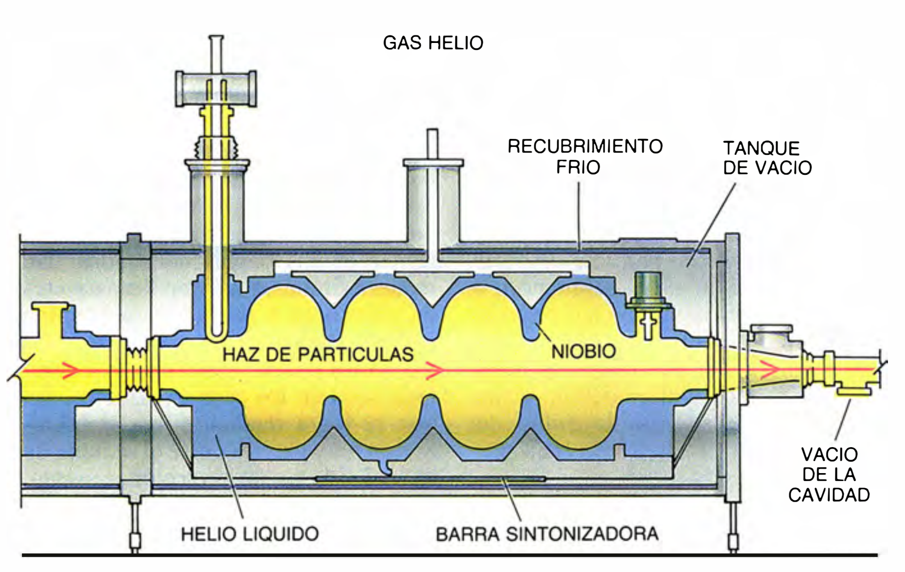
Tres máquinas perforadoras avanzaban a través de la roca a una velo-



5. LA CAVIDAD DE RADIOFRECUENCIA (RF) genera un campo eléctrico oscilante (*flechas rojas*) para acelerar los paquetes de partículas. A fin de reducir la pérdida de calor en las paredes de cobre de la cavidad, el campo de RF pasa a una cavidad esférica adjunta, de baja pérdida, cuando no hay paquetes para acelerar. La cavidad esférica disipa menos calor porque el gradiente de campo es máximo en el centro vacío de la esfera, en vez de serlo en su superficie metálica. Los gráficos ilustrados a la derecha indican las intensidades del campo en relación a los paquetes.

cidad de unos 25 metros diarios. Procedían de suerte que trabajaran con precisión centimétrica en la trayectoria deseada. Los componentes del colisionador que se hallan dentro del túnel se alinearon con una precisión de 0,1 milímetros. Esa justeza se logró gracias a un sistema de interferometría láser colocado en las colinas que rodean el lugar.

Para ahorrar tiempo y dinero, los ingenieros decidieron aprovechar los aceleradores menores ya existentes para el sistema de inyección de partículas, que las acelera hasta una energía elevada antes de inyectarlas en el LEP. Dos aceleradores lineales llevan las partículas primero a 200 MeV y después a 600 MeV. Un anillo de almacenamiento de 600 MeV acumula paquetes de partículas. El sincrotrón de protones del CERN se encarga de acelerar los paquetes hasta 3,5 GeV. Estos paquetes se inyectan luego en el supersincrotrón de protones; aquí se aceleran hasta 20 GeV antes de inyectarlos en el LEP.



6. UNA CAVIDAD SUPERCONDUCTORA DE RF, proyectada para elevar la energía del LEP a 100 GeV por haz. Las paredes de niobio refrigeradas por helio líquido reducirán las pérdidas de calor a cero. Los bordes, redondeados, eliminan los gradientes de campo agudos. Si no fuera así, los electrones se acumularían y calentarían el niobio por encima de la temperatura de superconducción.

Una vez que las partículas están dentro del anillo, hay que guiarlas por la trayectoria y focalizarlas en los puntos de colisión. De ese trabajo se encarga un sistema de electroimanes y placas electrostáticas. Cada uno de los ocho arcos, que representan más de las tres cuartas partes de la circunferencia del LEP, está formado por 31 módulos normalizados. Cada módulo mide 79,11 metros de longitud y contiene imanes en el siguiente orden: un cuadrupolo focalizador vertical (imán de cuatro polos que comprime el haz en su dimensión vertical), un corrector de órbita vertical, un grupo de seis dipolos curvadores, un sextupolo focalizador horizontal, un cuadrupolo focalizador horizontal, un corrector de órbita horizontal, un segundo grupo de seis dipolos curvadores y, por último, un sextupolo focalizador vertical.

Los imanes dipolares curvadores conducen a los electrones y positrones a lo largo de los arcos. Para reducir la radiación de sincrotrón, presentan un radio de curvatura grande, y así el campo magnético de estos dipolos es insólitamente bajo, en torno a 0,1 tesla. El campo bajo permite que el núcleo del imán dipolar tenga un diseño novedoso; en él, los intervalos entre las placas de acero del imán se llenan de cemento, como si fuese crema entre las capas de hojaldre de un milhojas. El modelo resultó ser un 40 por ciento más barato que los núcleos de acero habituales.

En el centro de cada punto de colisión, los haces de partículas deben comprimirse hasta dimensiones mínimas para aumentar la luminosidad del haz. Esta compresión se logra mediante un conjunto de imanes cuadrupolares superconductores, que focalizan el diámetro del haz a unas 10 micras en el plano vertical y a unas 250 micras en el horizontal.

Un enorme imán solenoide superconductor rodea el punto de colisión y genera un campo magnético uniforme en los detectores. El campo curva las trayectorias de las partículas cargadas que surgen de las colisiones y permite así a los experimentadores deducir su masa y su carga eléctrica.

Mediante el control de la corriente que pasa por sus bobinas se ajustan cuidadosamente todos los imanes del LEP. Se cuenta para ello con más de 750 fuentes de alimentación de corriente continua, que van desde menos de un kilowatt hasta siete megawatt. La precisión de las fuentes de alimentación llega hasta las dos partes en 100.000. Se sincronizan durante la etapa crítica en que los haces "suben" a la máxima energía.

En el LEP circulan, en sentido opuesto y al mismo tiempo, cuatro paquetes de electrones y otros cuatro de positrones igualmente espaciados. Los haces antagónicos cruzan sus trayectorias en el centro de cada detector y también en otros cuatro puntos intermedios. Sin embargo, durante la inyección, la acumulación y la subida de energía no se puede permitir que los haces contrarrotantes se acerquen entre sí, ya que los campos electromagnéticos asociados a cada paquete desviarían las partículas del paquete opuesto, haciendo que los paquetes explotaran contra las paredes del tubo del acelerador. Para superar este "efecto haz-haz", se los somete a un campo electrostático, que actúa como un separador de carriles invisible y mantiene separados los haces contrarrotantes en los ocho puntos de colisión posibles. Justo antes de que se esté preparado para registrar los datos, se desconectan las placas electrostáticas en los puntos de colisión; se deja que los haces choquen. (Los separadores pueden utilizarse también para provocar la colisión entre haces con mayor precisión.)

Como se ha mencionado antes, los electrones y positrones que circulan por el anillo de almacenamiento pierden continuamente energía en forma de radiación de sincrotrón. El papel de las cavidades resonantes de radiofrecuencia (sistema RF) es suministrar energía de repuesto a los haces.

El actual sistema de aceleración de RF consta de 128 cavidades de cobre de cinco celdas alimentadas por 16 klistrones de un megawatt. Los klistrones generan energía de radiofrecuencia que se manda a las cavidades donde produce un campo eléctrico intenso. El campo oscila justo a la frecuencia correcta para que, cuando un paquete de partículas entre en una cavidad, sea acelerado hacia adelante por el campo, a modo del practicante de surf a lomos de una ola perfecta.

El sistema de RF plantea un problema serio: los campos eléctricos acelerantes calientan las paredes de cobre de la cavidad y disipan, por tanto, una gran cantidad de energía. Para contrarrestar ese inconveniente, los ingenieros del LEP montaron una cavidad esférica encima de cada cavidad

acelerante, de suerte que la energía de RF oscilara entre las dos cavidades. La cavidad esférica está diseñada para que disipe mucha menos energía que la cavidad acelerante. Las oscilaciones se ajustan de manera que la energía esté en un máximo en las cavidades acelerantes en el instante en que la atraviesa un paquete de partículas. Durante la mitad del tiempo, la energía se halla en la cavidad esférica. Por tanto, los paquetes reciben la máxima aceleración, pero la pérdida por calor se reduce mucho debido a que la energía de RF pasa la mitad del tiempo en las cavidades de baja pérdida.

El sistema de aceleración de RF desempeña otra misión no menos importante: concentra las partículas en paquetes compactos. El sistema de RF opera a una frecuencia de 352,21 megahertz, lo que significa que genera 31.320 oscilaciones en una vuelta del LEP. Cada oscilación contiene una región donde cada partícula realiza oscilaciones estables con respecto a la partícula que está en el centro del paquete. Esta región estable, el cesto de RF, como se le llama, goza de propiedades que le permiten concentrar las partículas en paquetes discretos de uno o dos centímetros de longitud, en los que todas las partículas tienen la misma energía, con una precisión de 0,1 por ciento.

Puesto que hay 31.320 oscilaciones en una vuelta del LEP, habrá 31.320 cestos de RF. Podemos comparar el anillo de almacenamiento con un par de norias gigantes, dotada cada una de 31.320 barcas. Las norias giran en sentido opuesto. Si dos amigos quieren subir en norias distintas y cruzarse en un punto determinado, se deben sentar en las barcas adecuadas. De manera similar, para que los paquetes de electrones y positrones colisionen en los centros de los detectores, las partículas deben inyectarse y acumularse sólo en los cestos correctos. Lo que se logra mediante una sincronización muy precisa entre el sistema de RF y el sistema de inyección.

Después de acumular suficientes partículas y subirlas a la energía adecuada, los haces del LEP circulan normalmente unas seis horas, lapso durante el cual colisionan una y otra vez en los detectores. A lo largo de ese tiempo, cada una de las 10^{12} partículas de los haces recorre la circunferencia del LEP más de 240 millones de veces, viajando en total unos 6500 millones de kilómetros, más que la distancia de la Tierra a Neptuno. Todo el tubo del acelerador debe estar vacío a presiones muy bajas, ya que, en el recorrido de un trayecto tan largo, hasta algunas moléculas de gas chocarían con el haz y lo estropearían.

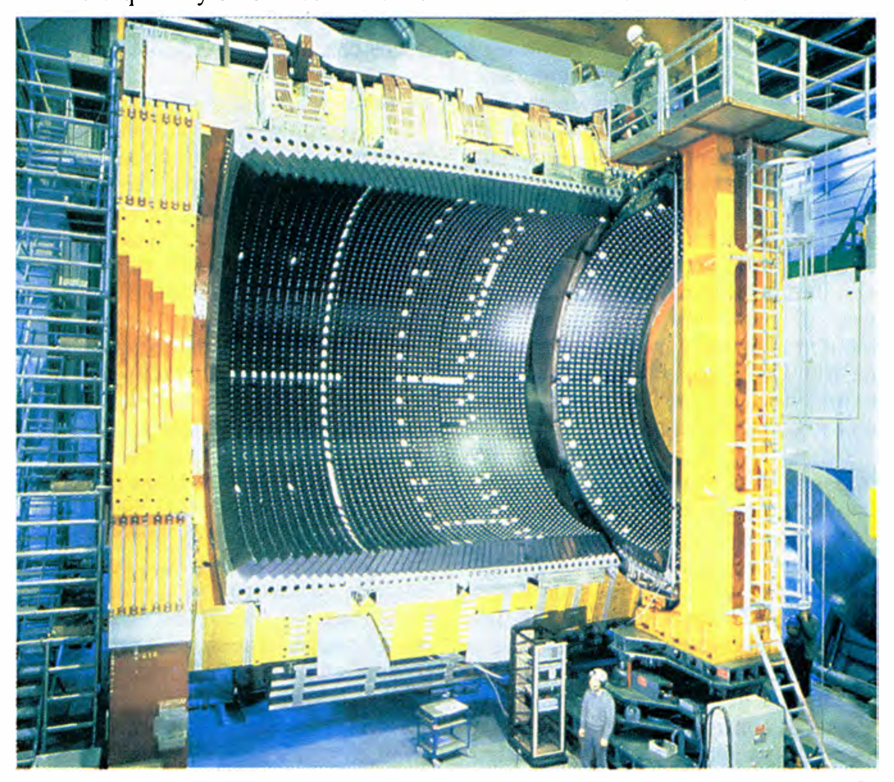
En el interior del tubo del acelerador del LEP la presión no alcanza los 10^{-11} torr. Ahora bien, hay 20 kilómetros, de los 27 de la cámara de vacío, que se hallan sujetos a la radiación de sincrotrón y ello provoca que las paredes de dicha cámara desprendan gas. Por tanto, en presencia del haz, la presión sube a 10^{-9} torr.

Sólo la mitad aproximadamente de la energía radiada se absorbe por las paredes de aluminio de la cámara de vacío. El resto la atravesaría y escaparía hacia el túnel, donde la elevada radiación destruiría rápidamente los materiales orgánicos (juntas de goma), cables y aislamiento eléctrico. La radiación podría también reaccionar con el aire, formando ozono y óxidos nítricos que, en condiciones húmedas, producirían ácido nítrico, muy corrosivo. Para evitar que la radiación escape, se cubre la cámara de aluminio con un revestimiento de plomo.

En anteriores anillos de almacenamiento de electrones, se conseguía el vacío ultraelevado mediante bombas lineales de desprendimiento ("sputtering") de iones, que deben funcionar en el campo de los imanes curvadores. En el LEP, los campos curvadores son demasiado bajos. Por eso incorpora un sistema nuevo, adoptado por primera vez en un acelerador, que se basa en bandas "captadoras no evaporables". Estas bandas, de tres centímetros de ancho, se extienden en conjunto a lo largo de 20 kilómetros. Consisten en un constantán (aleación de cobre y níquel) cubierto con una aleación de zirconio y aluminio. El material actúa a modo de tira matamoscas molecular al formar compuestos químicos estables con los gases más activos; las moléculas residuales de gas se limitan a "pegarse" a la tira.

Por último, para alcanzar presiones de menos de 10^{-11} torr, las superficies interiores de la cámara de vacío del LEP tienen que estar inmaculadas. Tras su fabricación, cada componente se limpió exhaustivamente con productos químicos y se almacenó en condiciones químicamente inertes. Tras su instalación en el túnel, todas las cámaras "se cocieron" a 150 grados Celsius durante 24 horas, bombeando agua supercalentada a una presión de ocho bar por los canales de refrigeración de las cámaras de aluminio y plomo. Válvulas, galgas, cámaras de acero inoxidable, separadores electrostáticos y otros componentes de ese tenor "se cocieron" también mediante calentamiento eléctrico.

Para que se produzcan las Z^0 que exige la comprobación precisa del Modelo Estándar los haces tienen que



7. LA PARED CURVADA de cristal plomado y hierro del detector Opal alcanza unos 10 metros de altura por 10 metros de longitud. El cristal plomado absorbe los electrones y los rayos gamma; gruesas placas de hierro pintadas de amarillo miden la energía depositada por los hadrones. Toda esta parte se une a una estructura idéntica para encerrar un gran imán solenoidal y cámaras llenas de gas que ponen de manifiesto las trayectorias seguidas por las partículas dotadas de carga.

ser muy luminosos. Se aumenta la luminosidad al maximizar el número de electrones y positrones en un paquete, y eso se consigue inyectando partículas de refresco en cada paquete circulante. La luminosidad se intensifica todavía más enfocando los paquetes por medio de una serie de lentes magnéticas, hasta convertirlos en un punto muy pequeño situado justo en el lugar de colisión. En su primera fase, el LEP proporcionará, así está proyectado, una luminosidad máxima de unas 2×10^{31} partículas por centímetro cuadrado y por segundo.

La consecución de luminosidad tan intensa ha supuesto un reto formidable. Un enjambre entero de fenómenos originados por la complicada física de los haces acelerados conspira para rebajar la luminosidad. El primero, el "campo de estela" inmediato. Los campos de estela se inducen en los componentes del acelerador por los campos electromagnéticos de las partículas que van al frente de un paquete. Los campos de estela excitan las partículas de la cola del paquete, aumentando por tanto la amplitud de sus oscilaciones transversales. El movimiento longitudinal inherente dentro de un cesto de RF envía estas partículas traseras a la parte frontal. Allí, debido a sus mayores oscilaciones transversales, las partículas generan campos de estela más intensos. Este crecimiento inestable de las oscilaciones lleva finalmente a las partículas contra las paredes del tubo del acelerador.

Podemos mitigar el efecto del campo de estela si creamos dicho tubo con perfecta simetría transversal y uniformidad longitudinal. En la práctica hay muchas irregularidades, que se presentan en regiones tales como las cavidades de RF, los separadores electrostáticos y los manguitos que enlazan las cámaras de vacío. Se ha dedicado mucho esfuerzo al diseño de estos componentes para minimizar el efecto del campo de estela. Los manguitos de vacío responden, en concreto, a un diseño completamente nuevo. Durante los primeros meses de funcionamiento realizamos mediciones del haz y descubrimos que el efecto del campo de estela del LEP se cifraba en un treinta por ciento menor de lo previsto.

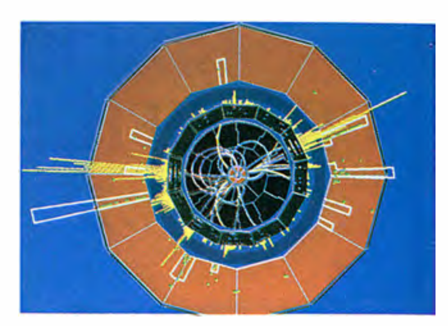
Entra dentro de lo posible que el bajo efecto del campo de estela aumente, en el futuro, la corriente del haz. Antes, sin embargo, habrá que prestar atención a otro problema. Hasta ahora, el LEP ha alcanzado una corriente máxima de 2,2 miliampère, un poco por debajo del objetivo, establecido en tres miliampère. De forma provisional, atribuimos esta reducida corriente a la magnetización del níquel que se empleó para curvar el revestimiento de plomo de la pared de aluminio de la cámara de vacío. De sólo 10 micras de espesor, el níquel crea campos magnéticos débiles, aunque dotados de intensidad suficiente para dispersar, bastante, el haz. Ese efecto se orilla con la introducción de campos magnéticos compensadores o con la desimantación del níquel.

El número de paquetes de partículas que circulan por el anillo de almacenamiento constituye otro importante factor que afecta la luminosidad. Ni que decir tiene que, cuantos más paquetes haya, más a menudo chocarán. Cuatro factores dejan sentir su peso en el número de paquetes. El primero es la potencia del haz. El aumento del número de paquetes incrementa la corriente del haz y, por tanto, su potencia. Durante la fase uno, la potencia del haz del LEP es de sólo 1,2 megawatt. Durante la fase dos, la potencia del haz será de 12 megawatt para la misma corriente de haz. Pero, si el número de paquetes crece, en un factor diez por ejemplo, la potencia del haz en la fase uno (55 GeV) sería de 12 megawatt y el mismo número de paquetes a 100 GeV consumiría 120 megawatt.

El segundo factor son los puntos de colisión extraños que deben eliminarse. Para un número dado de paquetes en circulación de cada tipo de partículas, hay el doble de puntos de colisión. A causa del efecto haz-haz, ya aludido, los haces deben mantenerse separados mediante un complicado sistema de placas electrostáticas en todos los puntos de colisión, excepto en los detectores. Si aumentamos el número de paquetes circulantes, deberemos instalar también más separadores.

El campo de estela remoto constituye el tercer factor. El campo de estela inmediato perturba la cola del paquete de partículas. El campo de estela remoto perturba los paquetes de partículas subsiguientes, lo mismo que la turbulencia de un avión a reacción zarandea a cualquier avión situado en su estela. La intensidad del efecto guarda relación directa con el espacio entre los paquetes. Se necesita ya un sistema de retroalimentación para corregir los efectos del campo de estela remoto generado por los cuatro paquetes de electrones y positrones que ahora circulan en el LEP. El coste y la complejidad de este sistema de retroalimentación crecerán en la medida en que se vayan introduciendo nuevos paquetes en el LEP.

Por último, debemos considerar la



8. PANTALLA DE ORDENADOR que representa un destello de partículas creadas por aniquilación electrón-positrón en el detector Aleph.

capacidad de los detectores para hacer frente a un ritmo mayor de colisiones. Se requiere un tiempo determinado para que los circuitos electrónicos de los detectores registren y midan un suceso. Un número mayor de paquetes aumentaría la tasa de sucesos y, por tanto, los detectores necesitarían un sistema electrónico más caro y complejo.

Por todas estas razones, la primera fase del LEP abarcaría, así se proyectó, sólo cuatro paquetes por haz. Si se dispusiese de más energía en la segunda fase, la máquina podría operar a las energías de la fase uno y cabría introducir más paquetes. Posibilidad que está en estudio. Exigiría la instalación de un nuevo sistema de separadores y de sistemas de retroalimentación del haz. Además, habría que modificar, de manera sustancial, la electrónica de los detectores. En el momento actual, parece viable un esquema con 36 paquetes por haz.

La primera fase del LEP se ha coronado con éxito en lo concerniente a la máquina propiamente dicha. Estamos intentando ahora aumentar la luminosidad hasta el valor proyectado, o quizá más allá. La segunda fase llevará los haces a energías de 100 GeV cada uno. Sin embargo, para alcanzar tales energías, el LEP precisará de mucha más potencia que permita reponer las pérdidas de radiación de sincrotrón. Más aún, con los sistemas de RF al uso, la potencia perdida por calentamiento se haría enorme. Resulta, por tanto, imprescindible menguar las pérdidas de calor en las paredes de la cavidad de aceleración hasta casi anularlas. Reducción que sólo se logrará con cavidades de RF superconductoras.

Hemos venido refinando el proyecto y la fabricación de cavidades superconductoras y esperamos incrementar el voltaje en ellas desde los actuales cinco megavolt por metro hasta siete megavolt por metro. En la fase dos del LEP se instalarán unas 200 cavidades de este tipo. También hay

que perfeccionar otros sistemas. En concreto, deberán sustituirse algunos cuadrupolos superconductores por otros más potentes y precisaremos nuevos klistrones para suministrar la potencia adicional. Si todo se desarrolla de acuerdo con lo previsto, los expertos del LEP podrán estudiar la física de las W^+ y W^- en 1994.

También estamos considerando un sistema de crear haces polarizados, formados por partículas con espines alineados longitudinalmente. Nos facultarán para abordar con mayor precisión las posibles desviaciones del Modelo Estándar.

Hay un grupo que estudia la posibilidad de construir un gran colisionador de protones en el interior del túnel del LEP. El Gran Colisionador de Hadrones (LHC, de "Large Hadron Collider") podría provocar la colisión entre protones a una energía de centro de masas de 16.000 GeV. Con semejante capacidad se situaría entre las energías del Tevatrón del Laboratorio Fermi y el Supercolisionador Superconductor. Otra opción sugestiva es la posibilidad de hacer chocar los haces de protones de 8000 GeV del LHC con los haces de electrones de 100 GeV del LEP.

No es fácil justificar a priori la decisión de construir el LEP, o cualquier otro nuevo acelerador. La bondad de la elección sólo puede valorarse retrospectivamente, tras unos diez años de funcionamiento. Los primeros meses del LEP han corroborado ya muchas de las razones esgrimidas para su construcción. Quizá la posibilidad más estimulante sea que dé lugar a descubrimientos no previstos y abra nuevos horizontes en el conocimiento de las leyes de la naturaleza.

BIBLIOGRAFIA COMPLEMENTARIA

ALEPH COLLABORATION: A PRECISE DETERMINATION OF THE NUMBER OF FAMILIES WITH LIGHT NEUTRINOS AND OF THE Z BOSON PARTIAL WIDTH. CERN-EP/89-169, 19 de diciembre de 1989. Enviado a *Physics Letters B*.

L3 COLLABORATION: MEASUREMENTS OF Z^0 DECAYS TO HADRONS AND A PRECISE DETERMINATION OF THE NUMBER OF NEUTRINO SPECIES. L3 Preprint n.º 004, 24 de diciembre de 1989.

OPAL COLLABORATION: A COMBINED ANALYSIS OF THE HADRONIC AND LEPTONIC DECAYS OF THE Z^0. CERN-EP/90-27, 23 de febrero de 1990. Enviado a *Physics Letters B*.

DELPHI COLLABORATION: A PRECISE MEASUREMENT OF THE Z RESONANCE PARAMETERS THROUGH ITS HADRONIC DECAYS. CERN-EP/90-32. Enviado a *Physics Letters B*.

Ciencia y sociedad

Días obscuros

Se resienten los pulmones, lloran los ojos. La piel sufre escozores con el contacto con gases corrosivos. Casi se agradece el sabor ocasional del azufre. Cuesta creer que en este ambiente pueda crecer algo. Pero no estamos describiendo ningún infierno imaginado por un ecologista apocalíptico; anotamos qué ocurre un día cualquiera de primavera en la región de extracción de lignito en Alemania oriental. Ciudades renombradas antaño por su fuerza cultural y política, Leipzig, Dresden o Weimar, son famosas hoy por la degradación de su entorno. Magdeburg, Cottbus, Plzen, Teplice, Cracovia, Katowice, Copsa Mica: la letanía sigue.

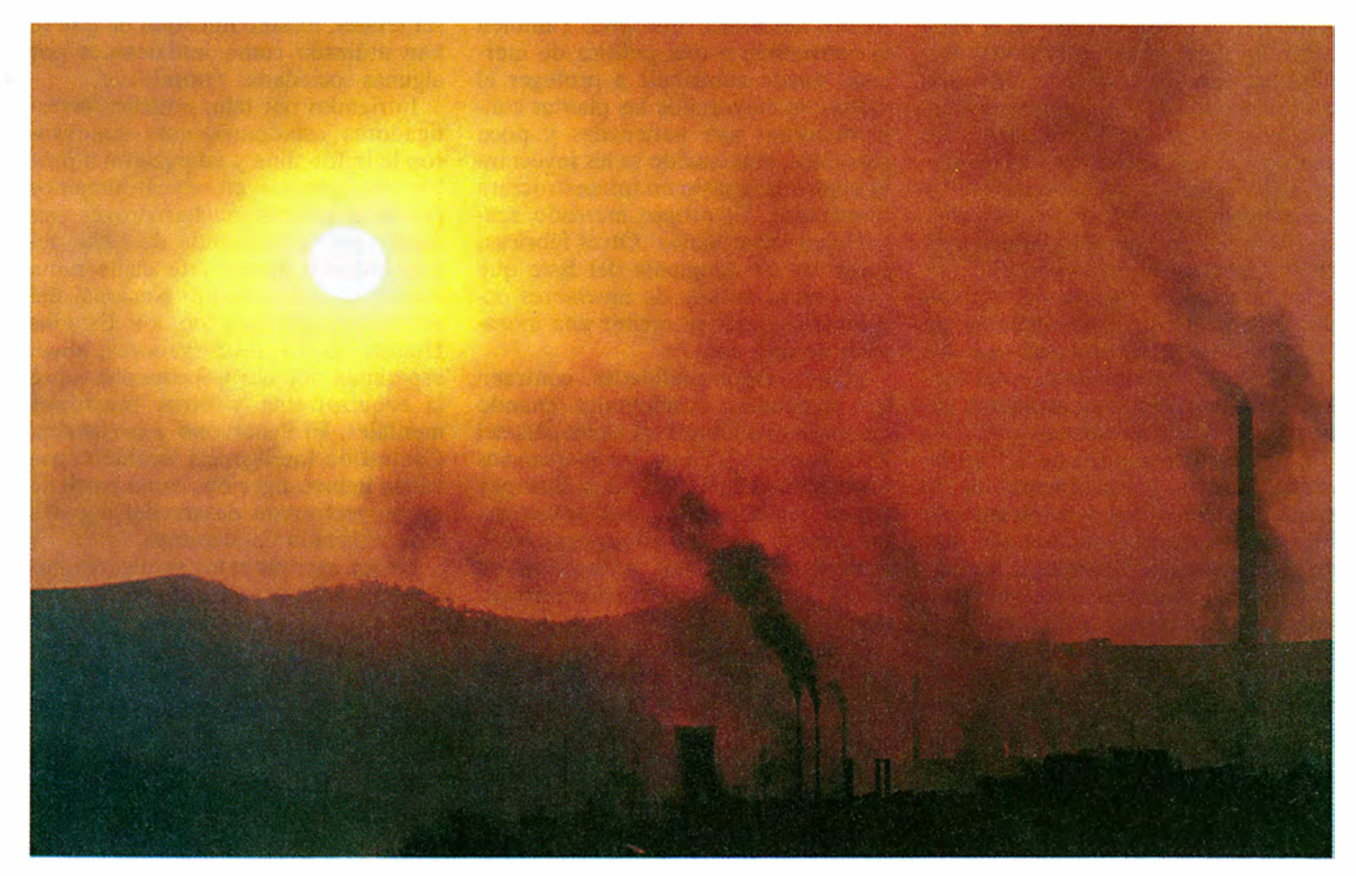
Estos lugares son viva muestra del peligro de una industrialización salvaje. Las estadísticas hablan de un recorte notable de la esperanza de vida; en el norte de Bohemia, las malformaciones congénitas alcanzan al diez por ciento de los niños, según los datos aportados por el partido verde checo. El apunte anecdótico es, también, contundente: la persona de mediana edad que limpia, a golpe de tos, sus pulmones en un área de descanso de la carretera en Teplice. A un cielo azul se interpone una cortina nubosa que sólo deja pasar un hilo triste y grisáceo de luz.

¿Presenta alguna característica singular la ruina ambiental de la Europa del Este? Para algunos, las lamentaciones vertidas por el destino de estas regiones velan el olvido de los países occidentales de sus propios desmanes en un pasado no demasiado lejano. A finales de los cuarenta, por ejemplo, la opacidad de la visión a mediodía en Pittsburgh significaba el progreso de la industrialización, y hasta comienzos de los sesenta la niebla letal se enseñoreó de Londres (niebla que provocó en 1953 la muerte de más de 2000 personas).

Antes del cambio de actitud, a finales de los setenta, la fundición Inco en la canadiense Sudbury arrojaba a la atmósfera el 1 por ciento del azufre emitido en todo el mundo cada año. No crecía un árbol en dieciséis kilómetros a la redonda.

La industria desarrollada bajo regímenes totalitarios no tiene por qué causar, en principio, más polución que la desarrollada bajo regímenes democráticos. A los ejecutivos occidentales les habría gustado poner entre rejas a los activistas que ensuciaban con lodo tóxico sus flamantes despachos, que detenían los vertidos de las chimeneas y promovían manifestaciones callejeras. Los gobiernos de los países del Este sí pudieron hacerlo. La información acerca del medio ambiente allende el telón de acero era materia reservada.

La planificación centralizada de los países socialistas tuvo su parte en la ruina. Pero los gobiernos occidentales también han encerrado a veces a ciertos ecologistas, por más que las compañías occidentales se hayan adaptado a los nuevos tiempos. Cuando la legislación ha quitado rentabilidad a la polución, las empresas han dejado de fomentarla. Bajo los regímenes socialistas no sólo se ha sufrido la destrucción del entorno, sino que, además, el proceso ha recibido la subvención estatal. En una economía



1. EL PUEBLO RUMANO de Copsa Mica se ha ganado a pulso la reputación de ser el lugar más contaminado del mundo. Foto de Gad Gross.

donde el dinero no importa, tampoco importan los costes de la polución. Las fábricas de Alemania oriental pagaron, en 1988, multas por contaminación de unos 30 millones de dólares; nadie exigía, es obvio, que estas fábricas rindieran beneficios.

En ningún sector se evidencia con mayor crudeza los incentivos del socialismo centralizado como en la industria energética: una energía cara podría haber estimulado la conciencia conservacionista, mas no en un país donde el estado paga la cuenta y presume del incremento anual de la producción de carbón. Alemania del Este, por ejemplo, consume cada año un 25 por ciento más de energía, a pesar de que su nivel de vida es mucho más bajo que el de sus homólogos occidentales. El aislamiento de las casas, la calefacción central y los dobles cristales en las ventanas son lujos desconocidos. Los habitantes de Praga derrochan, cada día, el doble de agua (a un sobrecargado sistema de alcantarillado) que los habitantes de Viena.

Ahora bien, el aire, azufrado y polvoriento, es sólo una parte de los excesos de la sociedad socialista. Países con pocos recursos energéticos (petróleo o gas natural), sus monedas carecen de la solidez suficiente para importar el petróleo que ha actuado como motor del desarrollo en el mercado libre. Se utiliza sobre todo lignito, un mediocre combustible que contiene casi un 90 por ciento de carbón y hasta un cinco por ciento de azufre y el resto silicatos y otros minerales. La economía centralizada ha acarreado otras catástrofes: en Checoslovaquia, por ejemplo, la tierra de labor se ha envenenado con altos niveles de cadmio a causa del empleo de abonos baratos procedentes de Gabón. (Sépase, sin embargo, que los fertilizantes, contaminadores del suelo y del agua, no son un problema exclusivo de países socialistas.)

La apretada geografía de la Europa central agrava los problemas de la contaminación. Ríos y acuíferos que reciben los vertidos industriales son los que abastecen de agua a las ciudades. El suelo arruinado es el mismo sobre el que la gente edifica sus casas y cultiva sus cosechas.

Con todo, ¿qué pueden hacer las naciones de libertad recién estrenada para lavar su culpa? Ahora que las fábricas ya no pertenecen al estado, funcionarios orientales y occidentales se lamentan de la falta de alguien que se responsabilice y obligue a instalar controles anticontaminantes. Sea o no del todo cierta esa queja, siguen abrigándose dudas sobre la capacidad de los gobiernos del Este para dictar nuevas leyes ambientales. Han de crear organismos competentes a partir casi de cero.

Y están por considerar los costes económicos. La descontaminación del aire, el agua y el suelo en Alemania del Este, por ejemplo, puede costar entre 40.000 y 150.000 millones de dólares. Naturalmente no toda la polución tiene su origen allí; durante los años setenta, más de 10 millones de toneladas de desechos tóxicos producidos por los occidentales se han depositado en los terrenos contiguos a la frontera divisoria. La nueva Alemania tendrá que buscar un nuevo emplazamiento para su continuo vertido de desechos.

Los checos aún no han acabado de analizar su medio ambiente. ¿Dónde están los problemas más graves? Por todas partes. El aire, el agua y el suelo. Jiří Nechvátal, de la comisión nacional del medio ambiente, se encoge de hombros. Por suerte, su país ha sido demasiado pobre y no ha comprado pesticidas. En Polonia, los funcionarios admiten que no tienen los 400 millones de dólares necesarios para detener la degradación ambiental, ni mucho menos los 25.000 millones de dólares que exige invertir la tendencia.

Quizá puedan recabarse algunos fondos del Banco Europeo. También la conversión a una política de mercado puede contribuir a proteger el medio: la mayoría de las plantas contaminadoras son anticuadas y poco rentables, aun cuando se les inyectara el dinero necesario en infraestructura limpiadora. El propio mercado acabará por desarmarlas. Otras fábricas, como las de Alemania del Este que están en la mirada de inversores occidentales, pueden recibir una moratoria de diez años.

¿Qué responsabilidades contraen las compañías occidentales cuando compran una fábrica en algún país del Este europeo? Todas las mercancías exportadas a la Comunidad Europea deben cumplir las normas ambientales, pero la Comunidad no regulará la calidad de las líneas de producción donde se han fabricado estas mercancías.

Tampoco perdurarán en el mercado libre los escasos hábitos de los habitantes del Este que producen beneficios ecológicos. Mientras que algún ministro occidental admira los sistemas de reciclaje de Alemania del Este que permiten aprovechar de nuevo el 40 por ciento del papel, vidrio y metales, otros censuran los altos precios pagados por estas prácticas.

Sapos y culebras

Rechoncho, amarillo verdoso y de textura guijosa, repugna a la vista. También puede ser una molestia, al envenenar a los perros y al provocar chasquidos bajo los neumáticos de los automóviles. Pero *Bufo marinus*, también conocido como sapo marino o sapo de la caña, se ha convertido en una celebridad internacional, despertando la histeria en la lucha antidroga en los Estados Unidos e instando transacciones comerciales en el Lejano Oriente. He aquí la historia, con toda clase de verrugas.

Este sapo grande vivía antaño tranquilamente en las regiones más cálidas de las dos Américas, llegando por el norte hasta California central. En la década de 1930 fue exportado a Australia en un intento de controlar los coleópteros que infestaban los campos de caña de azúcar (de ahí uno de sus nombres comunes). Como muchos anfibios, *B. marinus* rechaza a los depredadores mediante la secreción de una sustancia pegajosa y tóxica a traves de ciertas glándulas de su piel. La secreción contiene bufotenina, compuesto que se parece al neurotransmisor serotonina y que también se encuentra en hongos venenosos y plantas. Aunque las sustancias que poseen bufotenina pueden ser letales, existen informes de que se han utilizado como intoxicantes por algunas sociedades "primitivas".

Intrigados por tales noticias, investigadores estadounidenses sintetizaron la bufotenina y empezaron a probar sus efectos en seres humanos (junto con otras muchas drogas psicoactivas) en la década de 1950, según Stephen Szara, jefe de la rama biomédica del Instituto Nacional del Abuso de Drogas, de los Estados Unidos. Estos investigadores, dice, esperaban obtener información sobre la esquizofrenia y otros trastornos mentales. El Pentágono y la Agencia Central de Inteligencia también apoyaban la investigación, como parte de un proyecto para desarrollar agentes para el lavado del cerebro.

En un experimento, se inyectó bufotenina a cuatro reclusos de una prisión de Ohio. Sufrieron entonces efectos alucinatorios "que recordaban los del LSD y la mescalina", así como náuseas y dolor en el pecho, según señalaron los científicos en el número de *Science* del 18 de mayo de 1956. Los investigadores también observaron que "la cara de los reclusos adquirió un tono purpúreo como cuando se diluye el color de una berenjena. Ante esos efectos secundarios, recuerda Szara, se acabaron los

estudios sobre la bufotenina... de una manera oficial".

La Administración para la Ejecución de la Droga (DEA) prohibió la bufotenina a finales de la década de 1960. Por ironía de las cosas, la acción de la DEA indujo a algunas personas a intentar lamer sapos vivos, según indica Darryl S. Inaba, de San Francisco, ciudad que se convirtió en epicentro del culto a la droga de los años sesenta. Pero estos aventureros enfermaron en vez de transportarse a mundos de fantasía, y lamer a los sapos no llegó a ser una práctica extendida, añade Inaba.

Sin embargo, durante los dos últimos años los periódicos se han llenado de noticias extravagantes sobre malos tratos a sapos marinos. En abril de 1988, *USA Today* señalaba que los "hippies" australianos estaban "abandonando las drogas tradicionales por los sapos de la caña, a los que hierven con el fin de obtener un cóctel pegajoso, potencialmente letal". Aunque las autoridades australianas han negado la historia, ésta preparó a los medios de comunicación para difundir otras. Pocos meses después Inaba dio una conferencia sobre drogas en la que mencionó (según él, para introducir un matiz cómico) la práctica extravagante de los años sesenta de lamer a los sapos. Muy pronto se multiplicaron las llamadas de periodistas de todo el mundo que querían información sobre "los hippies enloquecidos que lamían sapos en las montañas", dice Alex Stalcup, director de clínica.

Inaba y Stalcup aseguraron a los periodistas que no había pruebas de que nadie estuviera ingiriendo secreciones de sapos, afirmación que confirmaron los funcionarios de la policía. Pero la historia persistía. "Los sapos reciben un lametón de drogatas desesperados", exclamó el *New York Post* el pasado enero. "¿Cuán bajo caerá la gente para intentar elevarse?", se mofaba el *Press Democrat*, de Santa Clara (California).

Inevitablemente, la realidad imitó a la ficción. El año pasado dos adolescentes de Nuevo México ingirieron toxina de sapo marino después de leer relatos sobre la "moda", y tuvieron que ser hospitalizados. Un joven australiano murió después de comer huevos de sapo marino. "El bulo ha causado mucha desgracia", dice Stalcup.

Mientras tanto, Australia se las tiene que ver con otros problemas relacionados. Los sapos marinos se han multiplicado tanto en la región nororiental del país que han expulsado a los anfibios indígenas y han envenenado a otros animales, silvestres y domésticos. Algunos perros se han convertido en empedernidos lamedores de sapos, según Glen Ingram, un herpetólogo del Museo de Queensland. "Les produce un efecto estimulante, quizá como el del alcohol", explicó a un periódico. Enfurecidos australianos han organizado excursiones de "erradicación de sapos"; palos de golf y bates de criquet, aunque se desaprueban oficialmente, son los instrumentos de ejecución favoritos.



2. BUFO MARINUS, *el sapo marino o de la caña, segrega veneno por las glándulas de la parte superior del dorso.*

Pero los ciudadanos de Brisbane pueden haber encontrado una perspectiva consoladora en la plétora de sapos. Recientemente persuadieron a China, cuya medicina tradicional incluye compuestos de otras especies de *Bufo*, para que explorara el potencial terapéutico del veneno de *B. marinus*. En abril, la Oficina de Desarrollo Económico de Brisbane despachó 100 gramos de esta sustancia a la Fundación Industrial de Shanghai. Otros brisbaneses están comercializando productos hechos con cuero de sapo. Aunque la piel de un solo sapo es algo pequeña, comenta Mark C. Underhill, de la oficina de desarrollo, se convierten en carteras, bolsos e incluso chaquetas "muy atractivas" si se cosen varias de ellas con destreza. Si no puedes lamerlos, únelos.

Metamorfosis del neutrino

Los resultados provisionales de un experimento realizado en la URSS parecen confirmar que el Sol produce muchos menos neutrinos de lo que las teorías actuales predicen. Si esa apreciación se confirma, habrá que revisar teorías del dominio de la física de partículas, de la astrofísica y de la cosmología. Según John Ellis, del Laboratorio Europeo de Física de Partículas (CERN), vivimos momentos de auténtica excitación.

Hay tres clases de neutrinos, correspondientes a las tres clases de leptones, a saber: electrón, muon y leptón tau. Detectados en 1956, los neutrinos son pieza esencial para los modelos de interacciones entre partículas y las hipótesis sobre evolución estelar.

En los años setenta los físicos comenzaron a sospechar que algo fallaba en estos modelos, cuando un detector emplazado en una mina en Dakota del Sur registró menos de una tercera parte de los neutrinos que se suponía procedían del Sol. (La colocación del detector en las profundidades evita señales falsas de otras clases de radiación cósmica.) Cuando a finales de los ochenta un detector de Japón corroboró el resultado, el déficit manifiesto de neutrinos solares se convirtió en uno de los mayores misterios de la física de partículas.

Sin embargo, algunos investigadores, advirtiendo que los detectores de Dakota del Sur y del Japón eran sensibles sólo a ciertos neutrinos de gran energía, han cuestionado la existencia de un déficit semejante en los neutrinos de baja energía, más abundantes.

Para resolver el problema, soviéticos y americanos han construido recientemente un detector de neutrinos poco energéticos que se ha instalado en el norte del Cáucaso. Denominado Experimento Soviético Americano Galio (SAGE) empezó a proporcionar datos en enero de este año.

El mes de junio, con ocasión del "Neutrino 90", congreso celebrado en el CERN, los investigadores hicieron públicos los resultados de los dos primeros meses de observación. El déficit de neutrinos observado por el SAGE supera incluso el observado con los otros detectores. De acuerdo con el modelo estándar de física solar, el detector del SAGE debiera haber registrado unos 14 neutrinos en un período de 60 días; no ha detectado ninguno.

Según Alexy A. Starobinsky, del Instituto de Física Teórica L. D. Landau en Moscú, si esta tendencia se confirma habrá que revisar de raíz las ideas vigentes sobre los neutrinos. Ya se habla de que los neutrinos electrónicos, la única clase detectable, se transforman antes de llegar a la tierra en neutrinos tau o muones. Este efecto lo predicen algunas teorías de la gran unificación, que intentan explicar con un formalismo unificado todas las fuerzas que rigen el comportamiento de la materia. Como hace notar Starobinsky, la profunda consecuencia de estas teorías es que al menos una clase de neutrinos tiene masa, quizá la suficiente para dar cuenta de la "materia oscura", u "oculta", que condiciona, en opinión de muchos cosmólogos, la evolución del universo.

Un reloj en los árboles

En 1970, cuando el arqueólogo John M. Coles empezó a excavar una senda de madera de 1800 metros de longitud enterrada en una turbera del sudoeste de Inglaterra, tenía la corazonada de que la estructura era antigua. Los análisis mediante radiocarbono confirmaron su sospecha. La madera databa de unos 4000 años a. de C., lo que significaba que el hallazgo era una reliquia del período Neolítico y era la carretera más antigua descubierta hasta entonces.

Sin embargo, Coles no se sentía del todo satisfecho, porque los resultados del método de radiocarbono tienen un margen de error bastante alto. Ahora Jennifer Hillam, de la Universidad de Sheffield, y Mike G. L. Baillie, de la Universidad de Queen, de Belfast, y sus colegas han dado una fecha mucho más precisa. La Pista de Sweet, así se la conoce, fue construida con árboles talados en el invierno de 3807-3806 a. de C.

Hillam y Baillie, que cuentan sus hallazgos en la revista *Antiquity*, encontraron la respuesta midiendo y contando los anillos de la madera. Cada anillo de un árbol representa un año de crecimiento, y el grosor de un anillo refleja las condiciones de crecimiento de aquel año. Por tanto, cuandos dos árboles muestran la misma secuencia de anillos delgados y gruesos, se puede presumir que los árboles han crecido bajo las mismas pautas y bajo las mismas condiciones ambientales.

El primer paso era construir una dendrocronología (cronología basada en los anillos de los árboles) maestra para la senda. Los dendrocronólogos hicieron diagramas complejos de secuencias a partir de muestras concretas de madera de la pista, y luego alinearon las pautas de un diagrama con las mismas pautas de las otras muestras, para producir una secuencia compuesta. Compararon más tarde esta cronología maestra con un segmento de una cronología datada de forma fiable, de unos 7000 años de edad, derivada de árboles que crecieron en Europa entre los años 5289 a. de C. y 1983 d. de C. La correspondencia indicaba que los árboles empleados para la Pista de Sweet crecieron entre 4202 y 3807 a. de C.

Un examen más minucioso de la madera mostró que los árboles fueron cortados después de formar anillos casi completos en 3807 a. de C., lo que significa que los constructores de la pista los talaron realmente después del otoño del año 3807 a. de C. y antes de que el crecimiento se reanudara en la primavera del año 3806 a. de C. Coles y su esposa, Bryony, que colaboraron en la excavación de la Pista de Sweet, suponen que los constructores procesaron la madera en los meses fríos, y luego construyeron la carretera, que atravesaba el pantano, a medida que el agua empezaba a descender en éste, al comienzo de la primavera de 3806 a. de C.

La cronología de la Pista de Sweet llena un vacío en una cronología más dilatada que en la actualidad se está desarrollando para Inglaterra. Asimismo, hace mucho más precisa la datación de reliquias neolíticas. Los Coles han demostrado que la carretera sirvió sólo durante unos 10 años antes de ser inundada por el pantano y enterrada en turba [véase "La carretera más antigua del mundo", de John M. Coles; INVESTIGACIÓN Y CIENCIA, enero de 1990]. Por ello es claro que cualesquiera artefactos encontrados junto a la estructura fueron hechos antes de 3806 a. de C. o dentro de una década, aproximadamente, después de esta fecha (por inferencia, estas fechas corresponden también a muchos útiles prehistóricos similares encontrados en otras zonas). Así, señala el matrimonio, fragmentos de cerámica abandonados junto a la pista indican que la tradición de la cerámica circular en la Inglaterra del Neolítico comenzó no más tarde que hacia 3796 a. de C.

Lo que es aún más sorprendente, un hacha de jade, procedente de los Alpes, prueba que los artículos producidos en el centro de Europa llegaban a Inglaterra hacia 3796 a. de C. "Una fecha que poca gente se habría creído antes", dice John Coles.



3. ANTIGUA CARRETERA, encontrada en una turbera en 1970, y que ahora está, literalmente, mostrando su edad.

Ciencia y empresa

Reciclar no es tan fácil

Los plásticos se crearon con la promesa de una vida mejor gracias al progreso de la técnica. Pero las palmas que les dieron la bienvenida se han vuelto lanzas, y ahora representan el papel de villanos de la película. Su incineración lleva consigo el riesgo de degradación de la atmósfera. Enterrarlos supone desperdiciar un suelo cada vez más escaso. En muchos lugares se aboga por la prohibición absoluta de envases de plástico. Por eso, los fabricantes de plásticos tratan de uncirse al carro de lo verde aireando un concepto ahora familiar: el reciclado.

La demanda de plásticos reciclados supera hoy, con mucho, la oferta. En principio, los plásticos pueden reciclarse hasta su forma original. Las materias plásticas constituyen el 7 por ciento en peso de la totalidad de los residuos municipales en los EE.UU. La cifra es comparable a los porcentajes de vidrio y acero y mucho mayor que la del aluminio. Esto se traduce en unos 13 millones de toneladas de residuos municipales de plástico por año. Las actividades industriales contribuyen con otros 8,6 millones.

San Hartwell, del Consejo de Defensa de los Recursos Naturales de la ciudad de Nueva York, pone en duda que se llegue a la sobresaturación "aun cuando se reciclara el 100 por cien".

Con gran diferencia, el plástico que más se recicla es el tereftalato de polietileno (TFP). Se usa en botellas de bebidas gaseosas; las fuentes industriales dicen reciclar un 12 por ciento del mismo. Se puede reciclar en una amplia variedad de objetos: recubrimiento de suelos, barreras señalizadoras de carreteras, fieltro de pelotas de tenis, cordeles, aislamientos térmicos e incluso repuestos para automóviles.

Para explotar este prometedor mercado, DuPont ha formado una empresa mixta con Waste Management, Inc., de Oak Brook, Illinois, para reciclar TFP y polietileno de alta densidad (PEAD, usado para jarras de leche). La empresa tiene dos fábricas en funcionamiento y prevé construir tres más para 1994. En conjunto, esas fábricas producirán 91.000 toneladas de plástico reciclado por año. Aunque indica un buen camino a seguir, se trata todavía de una gota en un balde comparado con los 0,7 millones de toneladas de TFP y 2,9 millones de toneladas de PEAD de desperdicios por año.

Más ambiciosos son los proyectos para iniciar el reciclado en gran escala de poliestireno, plástico común en la espuma de empaquetar. Aproximadamente dos millones de toneladas de poliestireno terminan en la basura cada año. Hasta ayer mismo, como quien dice, la tasa de reciclado de poliestireno era nula. Pero no hay que lanzar las campanas al vuelo. Pese a la sensibilidad ambiental y al interés económico, sólo se procesa un desmoralizador 1 por ciento. La industria del poliestireno asegura, empero, que la situación está a punto de cambiar.

En 1989, siete empresas −entre ellas Amoco Chemical, Dow Chemical y Mobil Chemical− formaron la National Polystyrene Recyclinc Company (NPRC), con la meta fijada en un 25 por ciento de tasa de reciclado para el poliestireno en 1995. Chris Vanden Heuvel, autorizado representante del sector, se muestra optimista porque el suministro de poliestireno se concentra en grandes restaurantes de comida rápida y en refectorios de instituciones. Así, por ejemplo, la Universidad de Harvard, el Colegio Universitario Boston, la Corporación Wang, McDonald's y bastantes hospitales y escuelas públicas están cooperando con el centro piloto de reciclado de NPRC en Leominster, Massachusetts.

Los fabricantes también se están moviendo. Rubbermaid vende productos de oficina hechos de poliestireno reciclado; Dolco Packaging, cartones de transporte de huevos, el primer uso de un producto reciclado para envasar alimentos. Artículos listos ya para el mercado son espuma aislante y cubos de basura de plástico.

Pero hay grandes obstáculos en el camino de un reciclado global. Los recipientes de poliestireno están frecuentemente manchados de residuos de alimentos, que pueden contaminar el plástico y crear emisiones durante el reciclado. Además, el volumen de los materiales tiende a encarecer el coste del transporte.

Por todo ello se teme que sea improbable el reciclado a gran escala del poliestireno. En opinión de Jan Beyea, científico de la Sociedad Nacional Audubon: "Es difícil imaginar que el reciclado llegue a alcanzar más del 5 por ciento por año. Pero la tasa de producción de residuos supera ese ritmo en muchos lugares". También observa que las compañías tropiezan con dificultades de inversión cuando abordan la infraestructura para el reciclado.

La fuerza que impulsa el reciclado de plásticos es, todavía, política y no económica. El efecto medioambiental de la prohibición de envases de plástico resulta discutible. (En Portland, la prohibición de envases de poliestireno para alimentos precocinados ha conducido al uso de cartón recubierto de polietileno, material no reciclable.) No obstante, la ola verde obliga a enfrentarse con el problema sin dilación.

Manipulación en la granja

Las cuatro terneras que pacen en el rancho texano de la fotografía tienen el mismo aspecto y comportamiento que cualquier bovino. Las apariencias engañan. Cada una porta genes de otras especies que promue-



VACAS A MEDIDA: las primeras vacas genéticamente manipuladas portan genes implantados que pueden adelantar su desarrollo y dar carne con menos grasa. Fotografía de Carter Smith.

ven un desarrollo más rápido y que su carne contenga menos grasa. Los genes fueron insertados en embriones por un grupo encabezado por Bert W. O'Malley, biólogo celular del Baylor College of Medicine.

Los fondos para la investigación procedían, sin embargo, de Granada BioSciences, ganaderos de Houston. Para Granada y un creciente número de empresas, tales animales genéticamente manipulados, o transgénicos, suponen futuros beneficios. Aunque ese tipo de ensayos se encuentran todavía en sus fases iniciales —rodeados, además, por la polémica—, un grupo de empresas pequeñas apuestan por la nueva ingeniería: Transgenix Sciences, DNX, GenPharm International y Embrex. Por su lado, DuPont, Merck, Genentech y Genzyme han acometido programas de transgénicos.

La manipulación genética de mamíferos encierra un potencial fecundo. Los ganaderos creen que la posibilidad de implantar nuevos genes en animales proporcionará una manera más rápida de lograr lo que tradicionalmente hacen los programas de mejora: esto es, crear animales con carne de mejor calidad y mayor resistencia a las enfermedades. Otras compañías están añadiendo genes humanos a animales de laboratorio para producir "modelos" de enfermedades humanas. Las cabras, conejos y ratones transgénicos sirven también de fábricas vivientes de medicamentos, produciendo proteínas de interés farmacológico en su leche.

La incidencia inmediata de los animales transgénicos se ha dejado sentir en los modos de investigación médica. En vez de buscar contrapartidas animales de enfermedades humanas, o de introducir tales condiciones mediante cirugía o fármacos, los investigadores en transgénicos diseñan animales modelo con los mismos genes que provocan el trastorno. Los ratones se han convertido en sujetos de estudio de muchos cánceres y virus, como el de la hepatitis. Se han desarrollado también ratones modelo para trastornos hereditarios, verbigracia, la distrofia muscular de Duchenne y la esclerosis múltiple.

La mayoría de las aplicaciones a corto plazo se darán en remedar enfermedades provocadas por un solo gen, o que al menos esté implicado; así, en el cáncer de mama o de colon. Pero aquello que la investigación clínica busca con mayor ahínco —modelos animales de la enfermedad de Alzheimer, el SIDA y otros males apremiantes— requerirá producir cambios de genes complejos y múltiples, lo que está hoy por hoy fuera de la mano de la ciencia. No se puede encargar un gen por teléfono; se trata todavía de un proceso de ensayo y error.

Para ayudar a los investigadores médicos en el diseño de animales de laboratorio transgénicos, el Instituto Nacional de la Salud (INS) otorgó en mayo un contrato de cinco años a DNX de Princeton, que le permite establecer y mantener en funcionamiento una instalación nacional de desarrollo transgénico. La compañía, que posee la patente sobre la técnica más usada para inyectar genes en embriones animales, facilitará que los científicos creen sus propios ratones transgénicos de investigación.

Cien investigadores pagarán al año cada uno 750 dólares por el trabajo de DNX para crear ratones transgénicos que porten sus genes únicos. El INS financiará el resto del coste real, de 7000 a 10.000 dólares por gen. "Hasta ahora, la gente había tenido que recurrir a los amigos y colegas de otras instituciones que tienen los medios", dice Joel M. Schindler, director de proyectos del Instituto Nacional de Salud Infantil y Desarrollo Humano. "Pensamos que era apropiado invertir dinero público en financiar una entidad para esta empresa común."

El gobierno participa también en un proyecto con ratones transgénicos para determinar la carcinogenicidad. El Instituto Nacional de Ciencias Ambientales de la Salud (NIEHS), de Research Triangle Park, Carolina del Norte, manipula ratones que llevan genes que se saben asociados a cánceres para comprobar su mayor sensibilidad a los carcinógenos que sus hermanos normales.

Los ratones en cuestión se prepararon en el laboratorio de Philip Leder, en la Universidad de Harvard; en abril de 1988 se convirtieron en las primeras formas de vida superior que recibieron una patente de los EE.UU. DuPont, que los cría bajo licencia de Harvard, los vende para investigación tumoral a unos 50 dólares cada uno.

Varias empresas de biotecnología han puesto a trabajar animales transgénicos en la producción de proteínas humanas que sirvan de fármacos. La mayoría de los esfuerzos se ha concentrado en genes que inducen que el tejido mamario segregue determinadas substancias en la leche del animal. La idea de convertir animales en "fábricas" de fármacos no es la de ser el primero en producir un medicamento, sino la de lograr hacerlo a un costo más bajo que las compañías que emplean bacterias o cultivos de células. Pharmaceutical Proteins de Edimburgo, Escocia, recurre a las ovejas para producir factores coagulantes que precisan los hemofílicos. Genzyme, de Boston, piensa emplear cabras para fabricar hormonas de crecimiento e interferón. Transgenic Sciences prefiere los conejos.

Los que levantan mayores objeciones por parte de las asociaciones para la defensa de los animales son los ensayos con ganado. Algunos de los primeros experimentos de inserción de genes en animales, realizados en la Universidad de Pennsylvania, dieron por resultado criaturas deplorables. Los descendientes de los primeros cerdos nacidos hace seis años con un gen para la síntesis de abundante hormona de crecimiento languidecen hoy en los establos de la estación de investigación del Departamento de Agricultura en Belstville. Su carne es más magra, como se esperaba, pero los cerdos sufren de artritis, letargia e inapetencia sexual.

El problema se puede resolver si se regula el gen, en vez de dejarlo actuar sin cesar. Los experimentos realizados por Transgenic con ratones abonan esa posibilidad.

Aviones solares

En tendido de sol, por favor, podría haber dicho Sanyo al encargar la zona para el proyectado despegue, el día 1 de julio, de su avión enano impulsado por energía solar. El vuelo desde San Diego hasta Kitty Hawk fue una gesta publicitaria para anunciar las nuevas pilas solares flexibles de la empresa. Sanyo proyecta incorporar esas pilas solares en toda clase de objetos de consumo, empezando por fuentes de energía portátiles que podrían estar a la venta en un año o dos. Las sombrillas de la playa, las tiendas de acampada, las velas e incluso los vestidos podrían suministrar energía, se asegura.

Las pilas solares se fabrican de silicio amorfo —no cristalino—, opción alternativa menos costosa, y menos eficaz, que las pilas cortadas de cristales de silicio. Hasta ahora, el revestimiento de silicio conversor de la luz se ha aplicado sobre substratos rígidos y bastante pesados: vidrio, acero inoxidable y plástico duro. Las pilas se han empleado, sobre todo, en relojes y calculadores manuales de bajo consumo. Pero, en el nuevo material, el silicio se aplica a ambas caras de una delgada lámina de plástico flexible, que se desliza después entre otras dos piezas de plástico para protegerlo. El resultado es una batería de pilas solares flexible y ligera.

Aunque Sanyo insiste en que su tecnología no infringe otras patentes, no andan otros tan seguros. El crédito

por los trabajos pioneros con silicio amorfo corresponde a Stanford R. Ovshinsky, fundador de Energy Conversion Devices (ECD). Esta y su subsidiaria Sovonics Solar Systems poseen 128 patentes sobre silicio amorfo y técnicas para hacerlo. Además, Sovonics ha venido desarrollando sus propias pilas solares flexibles, aproximadamente del espesor de dos bolsas de basura de plástico, durante cuatro años. La compañía afirma hallarse interesada en examinar lo que Sanyo ofrece.

Pero la destreza comercial de Sovonics no está a la altura de su experiencia técnica. Sovonics todavía no conoce beneficios, ni su empresa madre tampoco; y eso que llevan 30 años en la lid. Además, su material flexible está lejos de poder llegar a la fase comercial. En vez de encarrilar el material hacia el mercado de consumo, Sovonics decidió ofrecerlo a las empresas aeroespaciales al elevado precio de 100 a 200 dólares por watt. Hasta ahora, Sovonics ha suministrado sólo un pedido de sus delgadas y doradas hojas realizado por el Laboratorio Nacional Lawrence Livermore. ECD confía en que su última empresa cooperativa, formada con la japonesa Canon Inc., el mismo día del vuelo de prueba de Sanyo, le ayudará a abrirse un hueco en el mercado.

El éxito requerirá que ambas compañías escojan las aplicaciones con cuidado. Las pilas de silicio amorfo no tienen buen rendimiento de conversión de luz solar en electricidad. De hecho, la exposición a la luz provoca que la eficacia del material disminuya en un 10 por ciento en la primera semana de uso y otro 5 por ciento en un mes. Sanyo dice que el rendimiento de las baterías flexibles de su avión es de un 4 por ciento, comparado con alrededor del 20 por ciento de las pilas de silicio cristalino. Sovonics afirma que el rendimiento de su superior producto es de 5,5 al 6 por ciento. De esta manera, aunque las nuevas baterías sean ligeras, se requiere una gran superficie para generar una potencia adecuada. El planeador de Sanyo, de 90 kilos, presenta una envergadura de 17,5 metros cubierta de baterías solares. Estas generan la potencia de 2000 watt necesaria para mantener el ligero avión en el aire a velocidades de crucero de entre 65 y 145 kilómetros por hora.

A los consumidores no les preocupa el rendimiento, insiste Byron L. Stafford, del Instituto de Investigación de Energía Solar: "Les interesa más la potencia. Si da suficiente, estupendo." Sanyo coincide en ello: están convencidos de que se pueden hallar suficientes aplicaciones para las que basta una superficie razonable de baterías, aunque sean de bajo rendimiento.

Otro obstáculo potencial en la comercialización lo opone la duración. "Nadie coloca las calculadoras bajo un temporal de lluvia, y cuando dejan de funcionar, se las tira y se compra otra. Pero si se adquiere un producto más caro, queremos que dure más. Eso exige una buena dosis de ingeniería", apostilla Stafford. Sanyo asegura que su producto funciona bien a temperaturas entre −40 y 100 grados Celsius, pero el producto es demasiado reciente para saber cómo resistirá el paso del tiempo.

Pese a ello, Sanyo ha lanzado su apuesta: la estrategia que popularizó su primera generación de pilas solares rígidas −introducir productos baratos en el mercado en grandes cantidades− dará resultado de nuevo. Si el alto volumen de producción permite bajar los precios, la compañía puede estar en condiciones de dedicarse a aplicaciones más exigentes. Las baterías solares de silicio amorfo podrían todavía tener un lugar al sol.

Origen de la diabetes

La causa de la diabetes mellitus dependiente de insulina reside en la respuesta autoinmune de las propias células productoras de la hormona. Se empieza ahora a entender por qué se desencadena y mantiene esa reacción

Mark A. Atkinson y Noel K. Maclaren

Tiempo atrás, las personas que padecían la hoy llamada diabetes de tipo I o insulino-dependiente morían al cabo de un año, más o menos, de serles diagnosticada la enfermedad. Debíase el fallecimiento a que el páncreas perdía su capacidad de producir insulina, hormona indispensable para el metabolismo normal. En 1921, el aislamiento de la hormona a partir de páncreas de animales posibilitó el tratamiento de la enfermedad; desde entonces, ha resultado imprescindible para la supervivencia de millones de diabéticos.

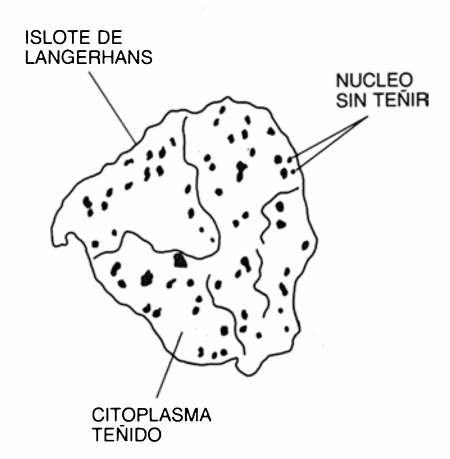
Ahora bien, ni la insulina animal ni la humana, de obtención más reciente, ofrecen la sanación completa de la enfermedad, y los pacientes deben inyectarse la hormona una o más veces al día a lo largo de toda su vida. Es más, muchos diabéticos sufren, con el tiempo, complicaciones de efectos devastadores. Por tratarse de una enfermedad crónica, con el tiempo se pueden dañar los vasos sanguíneos y producirse cardiopatías, parálisis, ceguera o disfunciones renales. Son frecuentes, también, las lesiones nerviosas.

Las refinadas técnicas de administración de insulina pueden frenar el desarrollo de tales complicaciones, pero la solución ideal es la prevención de la propia diabetes. Mas, para conseguirlo, tendrán los investigadores que descubrir las causas de la enfermedad. El estudio de la misma ha progresado, recientemente, a grandes pasos. Sabemos ya que la diabetes insulino-dependiente se origina por una respuesta autoinmune en el páncreas. Docenas de laboratorios de los Estados Unidos —incluido el de los autores en la Universidad de Florida— y del resto del mundo se afanan por sacar a la luz los pormenores de este proceso autoinmune. Nosotros en particular nos proponemos determinar qué componentes del sistema inmune son los principales agentes de esa respuesta, que desencadena la reacción y la mantiene.

Apoyados en ese trabajo, los médicos deberán, muy pronto, poder identificar el sujeto diabético insulino-dependiente, uno de cada 300, e iniciar así el tratamiento desde el primer síntoma. Se espera que, al doblar el siglo, puédase ofrecer una terapia preventiva sin riesgos.

El proceso autoinmune que causa la diabetes insulino-dependiente, altamente selectivo, suele comenzar antes de la edad adulta (por cuyo motivo se la llamó, en un principio, diabetes juvenil). El ataque no afecta a la mayoría de las células pancreáticas, que segregan enzimas digestivas, sino que se restringe a las células productoras de hormonas. Estas células se encuentran en agrupamientos periféricos —los islotes de Langerhans— que están dispersos por todo el páncreas. Incluso en los islotes, tres de los cuatro tipos de células se salvan del ataque; sólo las células β productoras de insulina, que ocupan la mayor parte del núcleo del islote, son víctimas de la reacción autoinmune.

La insulina facilita que la mayoría de las células del cuerpo capten los combustibles biológicos, glucosa incluida. Así, cuando las células β mueren y el páncreas cesa de producir la hormona, se acumula glucosa en la sangre, alcanzándose niveles anormalmente altos de dicho azúcar, síntoma distintivo de la diabetes. Entonces, el cuerpo se deshidrata, ya que los riñones trabajan a destajo para filtrar el exceso de glucosa hacia la orina. Mientras tanto, las células del cuerpo ayunan en medio de un mar de abundancia, por lo que degradan sin control sus reservas de lípidos y proteínas para proveerse de más energía. En virtud de este continuo catabolismo de los lípidos, se forman unos derivados ácidos: los cuerpos cetónicos. La presencia de éstos, combinada con la deshidratación, puede inducir coma y acarrear la muerte.



1. EN LOS ISLOTES DE LANGERHANS se alojan las células productoras de hormonas pancreáticas; entre ellas, las β que sintetizan insulina. Este islote de un corte de un páncreas normal se hizo visible por medio de anticuerpos fluorescentes, que se unen a un componente del citoplasma de la célula del islote (*verde*). Los anticuerpos procedían de la sangre de un diabético insulino-dependiente; reconocieron el islote sano, prueba de que el ataque autoinmune contra tejidos normales contribuye al desarrollo de la diabetes. La enfermedad se manifiesta cuando la mayoría de las células β se han destruido.

MARK A. ATKINSON y NOEL K. MACLAREN enseñan en el departamento de patología y clínica en la facultad de medicina de la Universidad de Florida en Gainesville. Atkinson inició su carrera en la Universidad de Michigan y se doctoró en la de Florida. Trabaja en proteínas pancreáticas implicadas en el desarrollo de la diabetes insulino-dependiente. Maclaren, jefe de departamento, se formó en Nueva Zelanda, en cuya Universidad de Otago se doctoró en 1963, especializándose en medicina interna en el Hospital Wellington. Estudió endocrinología pediátrica en la Universidad Johns Hopkins, antes de trasladarse a Florida en 1978 para dedicarse a la investigación sobre la genética y patogénesis de las enfermedades autoinmunes del sistema endocrino.

Las inyecciones de insulina interrumpen esta secuencia letal, e incluso la previenen, pero no pueden imitar el modelo normal de segregación de insulina del propio páncreas, ni tampoco normalizar el buen funcionamiento metabólico para prevenir las complicaciones a largo plazo de la diabetes, que se cree que derivan de la hiperglucemia sanguínea crónica, o se acentúan con ella.

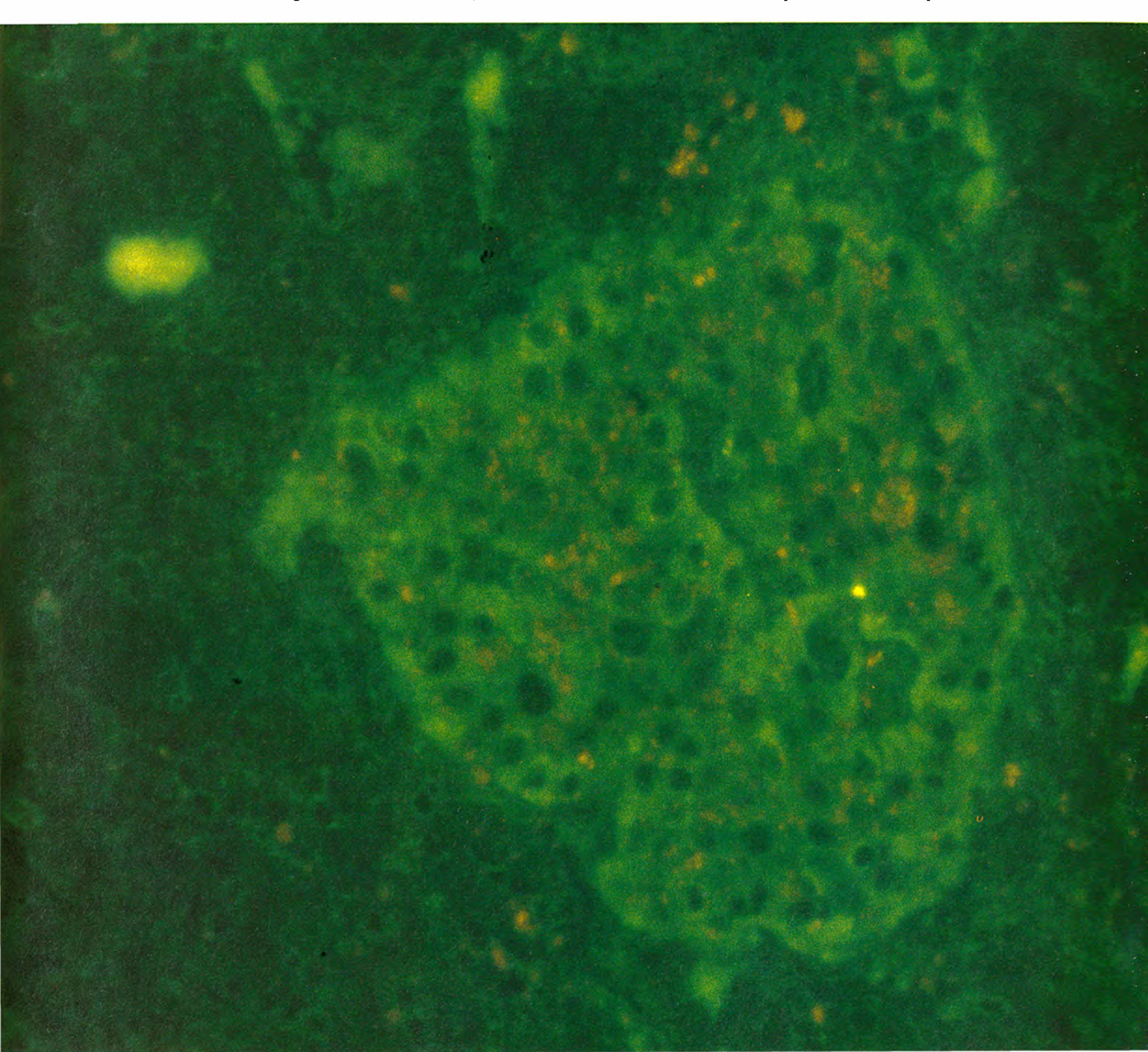
Pensábase, en un principio, que las células β se destruían de repente por acción vírica o quizá por la ingestión de algún producto químico tóxico. Esta hipótesis se basaba en la observación de los síntomas, que solían aparecer bruscamente. Se presentaban primero los principales signos de la enfermedad: apetito voraz (resultado de la necesidad energética de los tejidos), micción frecuente (debido al excesivo trabajo del riñón) y sed insaciable (causada por la necesidad corporal de reponer el agua perdida por la orina).

Los diabéticos pueden también perder peso con una rapidez asombrosa; y no es extraño que adelgacen hasta cerca de siete kilogramos en dos semanas. Algunos entran rápidamente en período de coma y, si no son sometidos de inmediato a una terapia adecuada, les sobreviene en seguida la muerte, aunque unos días o semanas antes pareciesen estar completamente sanos.

Por contra, los síntomas de la diabetes que no es dependiente de insulina (DMNID) o diabetes de tipo II —una enfermedad causada por mecanismos distintos (objeto todavía de estudio)— son más sutiles. Los diabéticos no-insulino-dependientes, generalmente mayores de 40 años y obesos en el momento del diagnóstico, producen diversas cantidades de insulina, pero no la aprovechan bien. Algunos de estos pacientes son tratados con insulina, para controlarles los niveles de glucosa en sangre, si bien no la requieren para su supervivencia diaria.

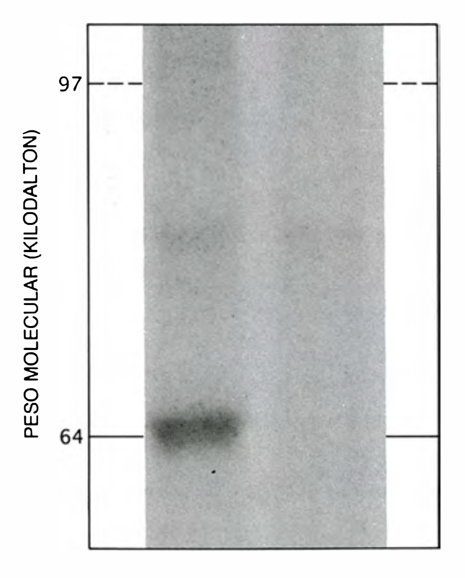
A pesar de las apariencias, la verdad es que el desarrollo de la diabetes insulino-dependiente no es tan rápido. Suele incubarse de manera silenciosa durante varios años, mientras el sistema inmune va destruyendo poco a poco las células β. Los síntomas clásicos se presentan sólo cuando se han destruido al menos el 80 por ciento de las células; el resto irán muriendo durante los dos o tres años siguientes.

El primer indicio importante de la



implicación del sistema inmune procede de los estudios de tejido pancreático de pacientes que murieron al poco tiempo de diagnosticárseles la enfermedad. Los exámenes de estas biopsias revelaban que los islotes de Langerhans se encontraban inflamados: estaban llenos de glóbulos blancos (linfocitos y monocitos), que suelen atacar a las células infectadas por algún germen y que, por otra parte, ayudan al cuerpo a luchar contra las infecciones. La insulitis, o inflamación de los islotes, podría explicarse, en teoría, como una defensa apropiada contra un organismo extraño. Esta respuesta podría darse si un microorganismo invadiera las células beta y presentase a continuación los antígenos —compuestos que señalan el invasor como "ajeno"— en la superficie celular.

Resultados posteriores apuntaban, sin embargo, a la autoinmunidad. Por ejemplo, muchos diabéticos insulinodependientes tenían otras enfermedades autoinmunes. Los investigadores detectaron también anticuerpos —o lo que puede llamarse "autoanticuerpos"— contra constituyentes naturales de las células de los islotes ("autoantígenos"), lo mismo en individuos de reciente diagnóstico que en los que estaban destinados a ser diabéticos años más tarde. Puesto que los anticuerpos participaban en la respuesta inmune como transductores de información (respuesta humoral), el descubrimiento de tales autoanticuerpos anti-islotes indicaba que el sistema inmune se hallaba, en realidad, activado contra sí mismo.



2. PROTEINA de 64 kilodalton (K) de peso. Aparece sólo en las células pancreáticas β e induce la producción de un anticuerpo específico contra la propia célula en los individuos diabéticos y pre-diabéticos, pero no en individuos sanos. Este "autoanticuerpo" se presenta antes que otros, lo que sugiere que dicha proteína 64-K ayuda a desencadenar la autoinmunidad que conduce a la diabetes. El autoanticuerpo se detecta por vía indirecta, al añadir a la sangre de distintos individuos proteínas de islotes marcadas con radiactividad. Los autoanticuerpos de la sangre precipitan cualquier proteína que reconozcan. La identidad de los autoanticuerpos se infiere de la separación de las proteínas (en función de su peso molecular) en un gel de electroforesis. Las muestras del gel que aparecen en la fotografía superior indican que la sangre de un diabético (*izquierda*) reacciona con la proteína 64-K (*banda oscura*), pero no lo hace con la sangre de un individuo sin diabetes (*derecha*).

El hecho de que se produzcan autoanticuerpos contra las células de los islotes no implica necesariamente que éstos sean los principales agentes de la destrucción de las células β, pero semejante posibilidad merece ser considerada. Los anticuerpos, producidos por linfocitos *B* activados (derivados de la médula ósea), se unen a un antígeno único. Cuando reconocen ese antígeno en una célula diana pueden bloquear directamente la propia función celular o bien pueden incrementar el poder destructivo de otros componentes del sistema inmune, como son los macrófagos (o células basurero), las denominadas células asesinas o "natural killer" (NK) y unas proteínas letales conocidas en conjunto como "el complemento". Estas proteínas se unen al extremo libre del anticuerpo, abriendo agujeros en la célula.

Se han estudiado con detalle tres autoanticuerpos, que se han encontrado en la sangre de los diabéticos y de sus familiares con mayor frecuencia que en la población normal. Uno de ellos, el autoanticuerpo contra el citoplasma de las células de los islotes (ICA), fue descubierto en los años setenta por Deborah Doniach, Gian Franco Bottazzo y sus colaboradores, en el Hospital Middlesex de Londres. Este anticuerpo reacciona con el citoplasma de todas las células de los islotes, reconociendo probablemente a un complejo lipídico denominado gangliósido.

Un decenio después, Steinunn Baekkeskov y Åke Lernmark, del laboratorio de investigación Hagedorn en Dinamarca, descubrieron un segundo autoanticuerpo. Se denomina autoanticuerpo 64-K, ya que reacciona con una proteína del islote de la que se sabe poco, salvo que su peso es de 64 kilodalton (K). Dicha proteína se encuentra sólo en la membrana plasmática de las células β.

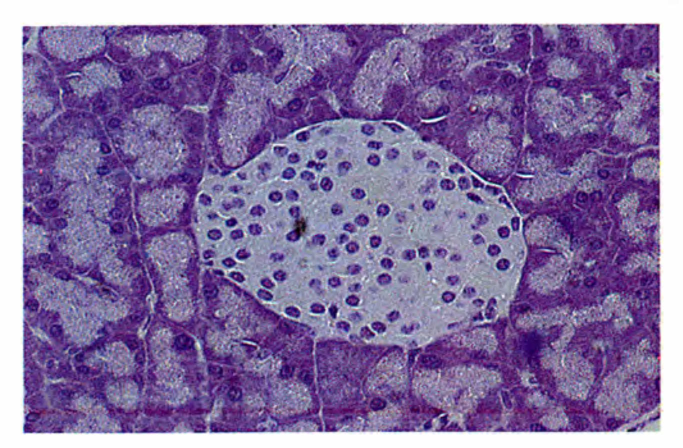
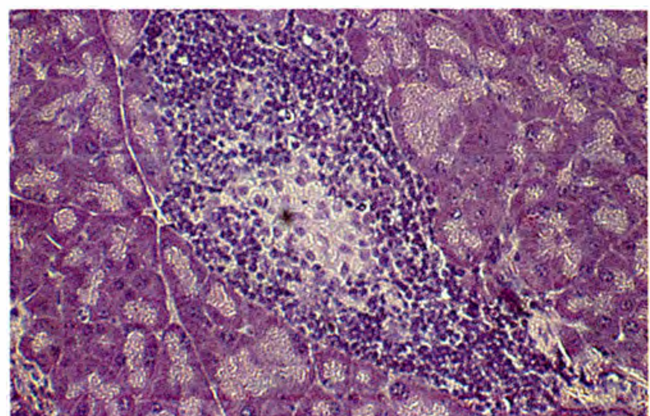
El tercer autoanticuerpo, descubierto por Jerry P. Palmer y sus colaboradores en la Universidad de Washington, se une específicamente a la insulina. Casi todos los pacientes tratados con insulina producen anticuerpos estándar que reconocen la hormona inyectada; sin embargo, podemos identificar estos autoanticuerpos en muchos diabéticos insulino-dependientes, antes incluso de la primera inyección de la hormona.

Si alguno de estos autoanticuerpos dirige su ataque contra las células β, se tratará probablemente de la variedad 64-K. Son anticuerpos específicos para dichas células (no como los ICA); aparecen antes que los otros dos tipos de autoanticuerpos y mucho antes de que la enfermedad pueda diagnosticarse. El tipo 64-K es también el único de los tres encontrado en los dos modelos de roedores diabéticos insulino-dependientes: la rata BB (del inglés "BioBreeding") y el ratón NOD (diabético no obeso); este último posee también autoanticuerpos contra insulina. En estas especies, el autoanticuerpo 64-K se produce a una edad temprana, coincidiendo con el comienzo de la insulitis.

Existen, empero, escasas pruebas de que este autoanticuerpo constituya el principal causante de la muerte de las células β pancreáticas. Si realmente los autoanticuerpos desempeñasen el papel principal en el desarrollo de la diabetes, su inyección en individuos sanos desencadenaría probablemente la enfermedad. Pero no se observa reacción cuando se administran a roedores sanos. De modo similar, no se sabe de ningún hijo de madre diabética insulino-dependiente que ya hubiese adquirido la enfermedad en el útero, aun cuando los anticuerpos maternos pasen directamente hacia el feto a través de la placenta.

La cuestión sigue abierta. No obstante, nosotros y muchos otros científicos sospechamos que la inmunidad mediada por células, más que por anticuerpos, es la principal responsable del desarrollo de la diabetes insulino-dependiente. En condiciones normales, la respuesta celular, protagonizada principalmente por linfocitos, es crucial para eliminar virus y otros agentes patógenos que infectan a las células.

El linfocito *T* coadyuvante ("helper"), que se forma en la médula ósea y madura en el timo, es la "autoridad principal", y posee receptores que reconocen a antígenos específicos. Los linfocitos *T* se activan sólo cuando el receptor se une al antígeno, que está a su vez unido a una molécula del CMH II en los macrófagos y en otras células presentadoras de antígenos. El CMH, o complejo mayor de histocompatibilidad, es una región del cromosoma 6 que codifica unas proteínas específicas de tejidos, o mar-

3. ISLOTE DE LANGERHANS de un ratón normal (*izquierda*) contrastado con el de un ratón NOD (diabético no-obeso), que es un modelo de diabetes insulino-dependiente (*derecha*); el islote de este último está lleno de linfocitos (*círculos oscuros*). Esta "insulitis" es señal de autoinmunidad.

cadores que distinguen los tejidos de un individuo de los de otro.

Una vez activadas, las células *T* coadyuvantes segregan citocinas (mediadores peptídicos) que incrementan la respuesta inmune. Una de estas citocinas, la interleucina-2, estimula la proliferación de los propios linfocitos *T* coadyuvantes, así como la de los citotóxicos, o células *T* asesinas. Los linfocitos *T* citotóxicos reconocen al antígeno en el caso exclusivo de que éste se halle unido a una molécula del CMH I. Otras interleucinas estimulan la secreción de anticuerpos por parte de los linfocitos *B*, incrementando, de este modo, la respuesta humoral.

Los resultados abonan la idea de que los linfocitos son los principales causantes de la muerte de las células *B* pancreáticas. Tanto en la rata BB como en el ratón NOD, la transferencia de células desde el bazo (rico en células NK y linfocitos *T*) de animales con una diabetes reciente hacia animales sanos puede inducir el desarrollo de la enfermedad en los segundos. De manera similar, Kevin J. Lafferty, del Centro de diabetes infantil Barbara Davis, en Denver, recopiló datos provisionales que sugieren que los linfocitos procedentes de sangre periférica de diabéticos insulino-dependientes, de diagnóstico reciente, pueden inducir insulitis en la llamada cepa de ratones desnudos, carente de linfocitos *T* normales.

En todo caso, no está claro si el mayor efecto en la destrucción de las células β pancreáticas en humanos es debido a los linfocitos *T* o a las NK, aunque nosotros nos inclinamos por los primeros. Aldo A. Rossini y Arthur A. Like, de la facultad de medicina de la Universidad de Massachusetts, y Alexander Rabinovitch, de la facultad de medicina de la Universidad de Alberta en Edmonton, han observado que las NK muestran mayor actividad durante el desarrollo de la diabetes en la rata BB. Pero otra línea de trabajo ha puesto de manifiesto que los linfocitos *T* son más activos en el ratón NOD, cuya diabetes se aproxima a la padecida por humanos. Los linfocitos *T* son la células que predominan en la insulitis humana.

¿Qué es exactamente lo que mueve al sistema inmune contra las propias células β pancreáticas del organismo? En un tiempo parecía que el desencadenante del proceso inmunitario residiera en una sustancia anormal presente en las células β; verbigracia, una proteína codificada por un gen mutante. Los descubrimientos recientes contradicen, sin embargo, esa explicación. Si un antígeno anormal participara en el desarrollo de la diabetes, los diabéticos de diagnóstico reciente o los prediabéticos generarían un autoanticuerpo contra dicho antígeno. Ahora bien, todos los autoanticuerpos de las células de los islotes de Langerhans identificados hasta la fecha reaccionan con alguno de los constituyentes normales de los islotes.

Por otro lado, nuestros colaboradores Ammon B. Peck y Drake M. LaFace han demostrado que, cuando el sistema inmune de los ratones NOD sanos (aún no diabéticos) se destruía y reconstituía con células de médula ósea procedentes de ratones normales, el ratón NOD no desarrollaba ni insulitis ni diabetes. Si las células β pancreáticas de los ratones receptores mostrasen un antígeno anormal, el sistema inmune normal lo reconocería y actuaría contra él.

A la inversa, cuando se sustituía el sistema inmunitario de ratones normales por células de la médula ósea de ratones NOD, los islotes de estos ratones receptores se inflamaban y muchos animales desarrollaban una diabetes manifiesta, aunque sus células β pancreáticas fueran normales al principio del experimento. Estos estudios arrojan también dudas sobre la vieja idea según la cual una infección vírica de las células β (que induciría el despliegue de antígenos foráneos) la causaba la diabetes insulino-dependiente.

La autoinmunidad de la diabetes podría provocarse por un proceso de "mimetismo molecular" [véase "Lo propio, lo ajeno y la autoinmunidad" por Irun R. Cohen; INVESTIGACIÓN Y CIENCIA, junio de 1988]. En tal proceso, un antígeno extraño —verbigracia, un constituyente de un virus o de otro microorganismo— podría instar una respuesta inmune normal en cualquier lugar del cuerpo. Si la conformación o la composición química de este antígeno corriese gemela a la de algún componente de la célula β pancreática, el antígeno podría también estimular un ataque inmunitario contra dicha célula β. El mimetismo molecular se halla claramente implicado en otras enfermedades autoinmunes, así en la fiebre reumática, que se puede producir después de una infección estreptocócica de la garganta.

Es probable, eso creemos, que la proteína 64-K sea una réplica de algún antígeno extraño, y provoque como tal el ataque autoinmune en las células β. Nos mueve a pensar así la aparición temprana de los anticuerpos contra esta proteína, que sugiere que la proteína 64-K adquiere importancia en los primeros estadios del ataque contra las células β.

Irun R. Cohen y sus colaboradores, del Instituto Weizmann de Ciencias de Israel, acaban de descubrir que ciertas proteínas bacterianas de "choque térmico", originadas durante el estrés, podrían contarse entre los antígenos extraños desencadenantes del ataque inmunitario. Cuando el grupo inmunizó ratones normales contra las proteínas e introdujo, en

ratones sanos, linfocitos *T* de aquellos animales las células transferidas indujeron insulitis y altos niveles de glucosa en sangre.

Podemos, pues, afirmar, a grandes rasgos, que una vez que algún agente disparador (¿una mímesis molecular?) induce una respuesta inmune (¿mediada por células?) contra las células β pancreáticas, el ataque autoinmune sigue el mismo curso que si dichas células productoras de insulina estuviesen infectadas por un virus. Los linfocitos *T* coadyuvantes dirigidos contra las células β se multiplican y reclutan otras fuerzas para la defensa del cuerpo.

Debemos, sin embargo, corregir este esquema elemental en dos aspectos. En primer lugar, las células β parecen acelerar su propia destrucción al evidenciar una insólita sensibilidad a la lesión mediada por el sistema inmune. En segundo lugar, el ataque inmune es más fuerte y prolongado que en la respuesta normal ante un virus.

Varias características de las células β nos explican la extrema vulnerabilidad de ésta. Dañadas, las células exponen un número excesivo de moléculas del CMH de clase I. Despliegue que aviva el ataque por parte de los linfocitos *T* citotóxicos. Las células β pueden mostrar, asimismo, particular sensibilidad a la lesión directa por citocinas. En este sentido, Jørn Nerup, del Hospital Steno de Copenhague, y Rabinovitch han demostrado que la interleucina-1, segregada por los macrófagos activados, resulta directamente tóxica para las células β, aunque lo es menos para otras células. Por último, las células β se manifiestan extraordinariamente expuestas a sufrir el ataque de radicales libres; los macrófagos y las células dañadas liberan esas moléculas de oxígeno altamente reactivas.

Somos bastantes los que sospechamos que la agresión inmune, desproporcionada, contra las células β posee más de una causa. Parece provenir, en parte, de una hiperreacción de los linfocitos *T* coadyuvantes ante los autoantígenos de las células β. Por otro lado, los supresores naturales del sistema, que mantienen a raya la actividad autoinmune, fallan en su cometido o lo desempeñan de forma apagada. Se conoce poco sobre el proceso normal de supresión de la actividad autoinmune; ignoramos, pues, por qué dichos procesos no interrumpen la destrucción autoinmune de las células β. Con todo, los estudios realizados con ratones NOD sugieren que el fallo podría tener base genética.

En estos animales, las anomalías genéticas en la región del CMH de clase II contribuyen a que sus portadores sean propensos a desarrollar diabetes. En el ratón, los genes de dicha región codifican dos tipos de proteínas. Una, llamada IE, parece intervenir de forma destacada en la actividad normal del supresor. La otra, IA, es primordial para el reconocimiento antigénico por los linfocitos *T*.

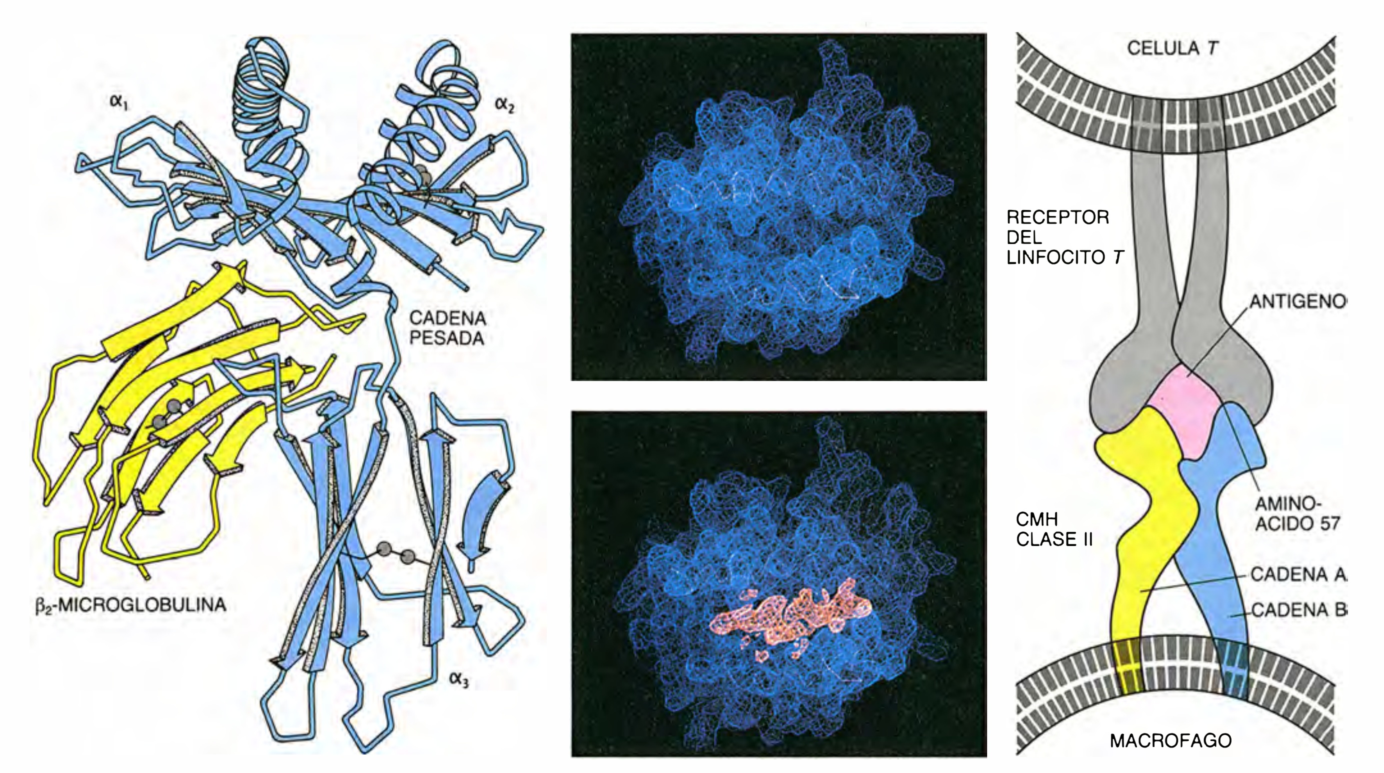
De ello se deduce que los ratones NOD no expresan el gen IE (y, por tanto, no sintetizan la proteína IE). Sin embargo, Hirofumi Nishimoto, de la Universidad de Osaka, demostró que, cuando en los ratones NOD se induce, por transferencia génica, la expresión del gen, éstos apenas desarrollan la insulitis, si es que lo hacen. En consecuencia, la ausencia del gen IE debe contribuir al desarrollo de la diabetes. Pero la carencia de dicho gen no constituye la única causa de la diabetes, ya que, si así fuera, tal ausencia provocaría la enfermedad en otras cepas de ratón. A pesar de ello, existen cepas de ratones normales que no expresan la proteína IE, ni desarrollan insulitis ni diabetes. De cualquier forma, no está claro si, en la diabetes humana, la autoinmunidad guarda relación con fallos en la expresión o en la estructura de algún gen análogo al IE.

Más fructífera ha resultado la línea de trabajo centrada en descubrir la razón de la desproporción de la respuesta autoinmune ante las células β. En el ratón, la herencia de los códigos genéticos de una variedad determinada de la proteína IA —la única de los ratones NOD— aumenta la proclividad a la diabetes.

Edward H. Leiter, del Laboratorio Jackson en Bar Harbor, ha demostrado que al menos otros dos genes (cuyas funciones no acaban de desentrañarse) contribuyen al desarrollo de la diabetes en los ratones NOD. Una vez se caractericen mejor estos genes, podrán servir de sondas para identificar genes similares en humanos.

La región CMH II se ha asociado también con la sensibilidad a la diabetes insulino-dependiente en humanos. Esta región D, así se la conoce, comprende tres loci —DP, DQ y DR—, cada uno de los cuales codifica una sola molécula que está formada por dos cadenas aminoacídicas: A y B. Hasta la fecha sólo se han relacionado con la diabetes los loci DQ y DR. Los genes de ambos loci poseen cierto número de variantes o alelos; por lo que existen muchas proteínas DQ y DR diferentes en la población

4. LOS GENES del complejo mayor de histocompatibilidad (CMH) que se encuentran localizados en el cromosoma 6 se han asociado con la diabetes. Los genes sospechosos de estar vinculados con la enfermedad se alojan en la región D, que está dividida en tres loci: DP, DQ y DR. Los tres genes codifican moléculas del llamado CMH II, todas ellas formadas por dos cadenas de aminoácidos (A y B). Ya que algunos de los genes que determinan estas cadenas son polimórficos, o variables (*en color*), existen diferentes tipos moleculares. Los genes de las proteínas DR1, DR3 y DR4 son frecuentes en los diabéticos insulino-dependientes. Es posible que no confieran directamente la sensibilidad a la diabetes, pero los genes DQ, a los que se encuentran ligados, sí lo hagan. Cuando en la cadena B de DQ aparece ácido aspártico en la posición 57 se dice que el gen correspondiente es "protector"; si aparece otro aminoácido en esa posición, aumenta la sensibilidad a la diabetes.



5. ANALISIS de la estructura de las moléculas del CMH. Permite comprender mejor por qué la carencia de ácido aspártico en la posición 57 de su cadena B de la proteína DQ del CMH II puede contribuir al desarrollo de la diabetes. Don C. Wiley y sus colaboradores, de la Universidad de Harvard, han visto que una molécula de la clase I (*izquierda*) consiste en una cadena α pesada (*azul*) y una menor, la β_2-microglobulina (*amarillo*), unida reversiblemente al polipéptido. Los dos primeros dominios de la cadena α forman una hendidura donde se va a unir el antígeno. Las representaciones de Wiley y sus colaboradores (*centro*) también muestran la hendidura vacía vista desde arriba (*parte superior*) y ocupada por un antígeno (*rosa, parte inferior*). Las moléculas de la clase II, que en los macrófagos presentan el antígeno a los linfocitos *T* coadyuvantes (*derecha*), son probablemente similares a las moléculas de la clase I, con dos salvedades: su cadena A se asemeja a una combinación del primer dominio α y la β_2-microglobulina, y la cadena B se parece al resto de la cadena α. Se cree que el aminoácido 57 de la cadena B se encuentra en la superficie de la hendidura de unión de la molécula de la clase II, donde podría afectar a la intensidad de unión del antígeno o al reconocimiento de éste por los linfocitos *T*. El ácido aspártico parece contribuir a la debilidad del enlace de los antígenos relacionados con la diabetes, por lo que en este caso no se produce respuesta autoinmune. Otros aminoácidos podrían engarzarse con vigor a dichos antígenos y, por tanto, desencadenar una respuesta intensa.

(todas ellas clasificadas). Se expresan los genes DQ y DR heredados de ambos progenitores (son codominantes); por tanto, cada individuo puede tener dos tipos de proteínas DQ y dos DR.

Hace ya tiempo se descubrió que los genes que determinan las proteínas DR1, DR3 y DR4 abundan en los diabéticos insulino-dependientes, mientras que escasean los que determinan DR2 y DR5. El hallazgo sugirió que los genes DR1, DR3 y DR4 incrementaban la proclividad a la diabetes o que su herencia estaba ligada a algún gen "sensibilizador", así como que DR2 y DR5 conferían cierta protección frente a la enfermedad o se heredaban junto con algún gen "protector".

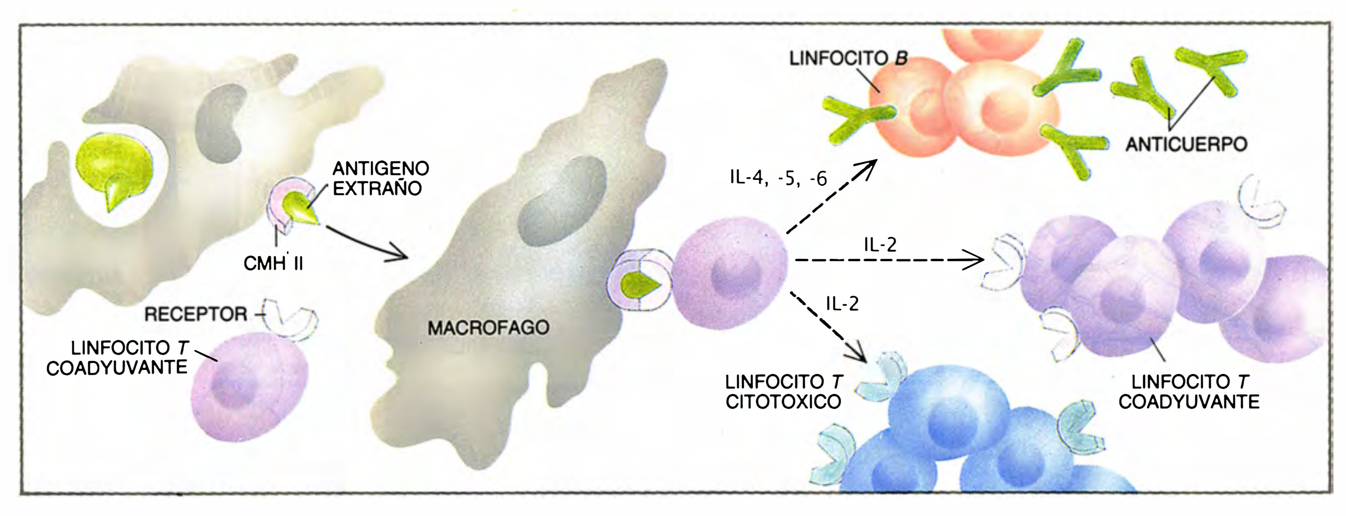
La investigación actual se dirige hacia genes sensibilizadores específicos. Por citar una muestra, se sabe que la herencia de la información genética correspondiente al gen DR4 suele ir ligada al gen que codifica para cierta proteína DQ: la DQw3.2 (la "w" constituye una designación provisional; el número decimal distingue a esta proteína de otras proteínas DQw3 que resultan semejantes en los tests serológicos, si bien poseen pequeñas diferencias en su secuencia de aminoácidos). Los individuos que heredan el código genético de la proteína DQw3.2 son bastante propensos a sufrir diabetes, mientras que los relativamente escasos individuos DR4-positivo que poseen el alelo DQw3.1 presentan menor probabilidad de desarrollar la enfermedad.

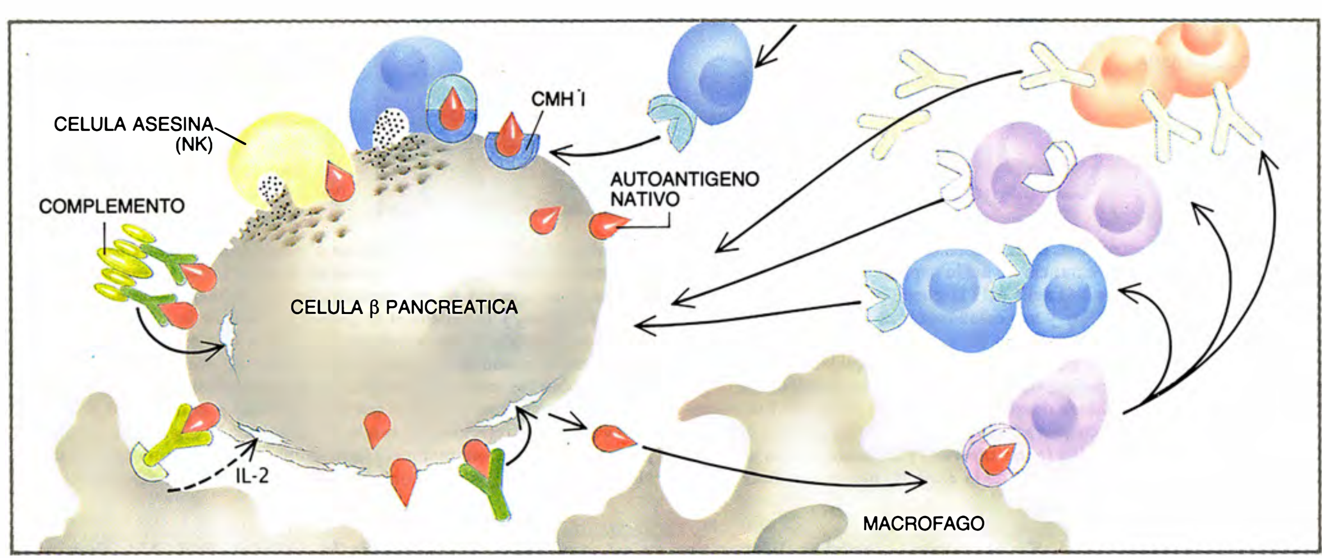
Hace unos dos años, John A. Todd y Hugh O. McDevitt, de la Universidad de Stanford, identificaron una región de moléculas DQ que incrementaba la proclividad hacia la diabetes. En los individuos que heredaban los genes DR1 a DR6 encontraron, en general, que la sensibilidad a la enfermedad venía determinada por el aminoácido 57 de la cadena B de la proteína DQ. Cuando esta posición estaba ocupada por el ácido aspártico, que posee carga negativa, la probabilidad de desarrollar la diabetes era baja. Sin embargo, cuando se encontraba un aminoácido apolar, como la valina o la serina (lo que ocurre en la proteína DQw3.2), aumenta el riesgo.

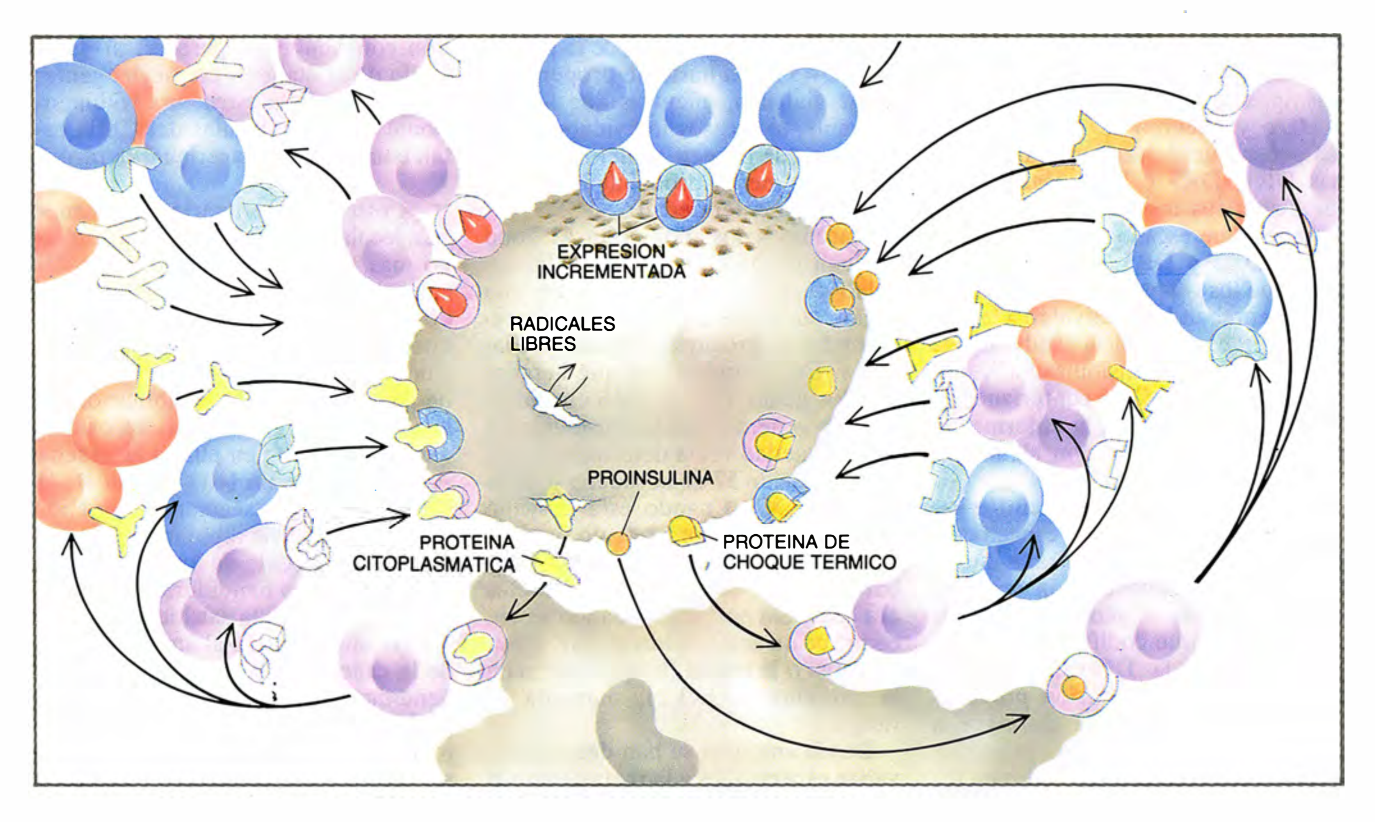
Desde entonces se han descubierto varias excepciones a la regla de que el ácido aspártico de esa posición protege de la enfermedad. No obstante, la carencia del aminoácido en cuestión contribuye, al parecer, al desarrollo de la diabetes en mucha gente; los análisis de la estructura tridimensional de las moléculas del CMH están comenzando a averiguar la razón.

La estructura primaria de las moléculas del CMH I se ha estudiado más a fondo que la de las moléculas del CMH II. A los cristalógrafos de rayos X debemos, sin embargo, una idea razonable de la conformación de las moléculas de la clase II. Se ha observado que la combinación de las cadenas A y B forma, aproximadamente, una X; en ella, a la cadena A corresponde un lado y, a la B, el otro. Juntos, los segmentos superiores de las cadenas dibujan una figura similar al panecillo de un "perrito caliente", en cuya hendidura se encuentra el antígeno ("la salchicha").

Este modelo sitúa al aminoácido 57 de la cadena B en la superficie de la hendidura formada por la parte superior de las dos cadenas, en un punto accesible tanto al antígeno como al receptor de los linfocitos *T*. Las propiedades del aminoácido de esa po-


LINFOCITO B
ANTICUERPO
ANTIGENO EXTRAÑO
IL-4, -5, -6
CMH II
IL-2
RECEPTOR
MACROFAGO
IL-2
LINFOCITO T COADYUVANTE
LINFOCITO T CITOTOXICO
LINFOCITO T COADYUVANTE


CMH I
CELULA ASESINA (NK)
AUTOANTIGENO NATIVO
COMPLEMENTO
CELULA β PANCREATICA
IL-2
MACROFAGO


EXPRESION INCREMENTADA
RADICALES LIBRES
PROINSULINA
PROTEINA CITOPLASMATICA
PROTEINA DE CHOQUE TERMICO

6. MODELO DE AGRESION CONTRA LAS CELULAS BETA

A. COMIENZA EL ATAQUE con una respuesta inmune vigorosa contra un antígeno ajeno (*verde*) que remeda a un componente normal de la célula β. Cuando los macrófagos han fagocitado a un invasor en algún punto del organismo y presentan el antígeno "gemelo" (trabado a una molécula del CMH II) ante los linfocitos *T* coadyuvantes, estos últimos segregan péptidos —interleucinas (IL)— que promueven la activación de otros linfocitos *T* coadyuvantes, linfocitos *T* citotóxicos y *B* productores de anticuerpos.

B. EN EL PANCREAS, los linfocitos *T* citotóxicos sensibilizados reconocen al antígeno natural (*rojo*) gemelo del antígeno extraño —un autoantígeno— en las células beta dondequiera que esté unido a moléculas del CMH I, moléculas ubicuas. Los anticuerpos se unen también a las células β, dañándolas directamente o bien solicitando ayuda de otros componentes del sistema inmune, como por ejemplo las proteínas tóxicas que constituyen el llamado complemento. Muy pronto, los macrófagos entrarán en escena, estimulando a los linfocitos *T* coadyuvantes para intensificar la destrucción mediada por células.

C. EN EL SIGUIENTE ATAQUE la célula β es dañada aún más. Dichas células sintetizan un exceso de moléculas del CMH I (con lo que promueven una mayor actividad de los linfocitos *T* citotóxicos). También muestran un gran número de moléculas de la clase II. Esta aumentada exhibición de moléculas del CMH inducirá la supresión de la autoinmunidad, pero en los prediabéticos ocurre lo contrario: estimula con vigor la actividad de los linfocitos *T* coadyuvantes. A medida que la célula beta se va dañando más y más, ciertas sustancias, que suelen hallarse en su interior (así, la proinsulina), comienzan a aparecer en la superficie de la célula. Estas moléculas producen nuevas oleadas de ataques, que terminarán destruyendo a las células beta.

sición pueden condicionar, y bastante, la conformación de la hendidura (como puede ser el ángulo que se subtiende entre las superficies de las cadenas A y B) y, de ese modo, influir en la fuerza con que se unen los diferentes antígenos. Por tanto, una hendidura con un aminoácido apolar en la posición 57 trabaría al autoantígeno responsable de la diabetes con especial vigor. Esta unión cerrada aumentaría la probabilidad de reconocimiento por los linfocitos *T*. Cabe también que el contacto directo entre el aminoácido 57 y el receptor de los linfocitos *T* intervenga en la actividad de dichas células *T*.

Ulteriores estudios estructurales, más amplios, podrían ayudar a explicar por qué la presencia del ácido aspártico en la posición 57 no protege de la diabetes a todos. Al haber varios puntos de contacto entre cualquier antígeno y la hendidura de unión, la carga o la configuración de un aminoácido particular en cualquiera de esos puntos podría influir en la fuerza con que las moléculas DQ presentaran el antígeno.

Es de esperar que se descubran también asociaciones entre diabetes humana y genes que se hallen fuera del complejo mayor de histocompatibilidad. Lo mismo que ocurre en los ratones NOD, muchos genes contribuirán, sin la menor duda, a la sensibilidad a la diabetes.

Los hallazgos de la genética y otros datos recopilados hasta la fecha pueden barajarse en distintos modelos relativos al desarrollo de la diabetes insulino-dependiente. De acuerdo con el nuestro, el individuo genéticamente predispuesto a desarrollar la enfermedad se ve expuesto a un antígeno foráneo, que remeda algún constituyente de las células β pancreáticas (quizá la proteína 64-K). A continuación, un macrófago u otra célula presentadora de antígenos expone dicho antígeno asociado a una molécula del CMH II, que se une al antígeno con fuerza inusitada (como en el caso de la proteína DQw3.2). El reconocimiento de dicho antígeno por un linfocito *T* coadyuvante desencadena una respuesta inmune, muy intensa y prolongada, contra el antígeno extraño, que puede ser avivada posteriormente si los receptores de los linfocitos *T* se unen vigorosamente a la unidad antígeno-CMH II.

Por fin, el torrente saguíneo transporta las células inmunes excitadas, así como los anticuerpos específicos, hasta el páncreas; aquí, las células β, algunas por lo menos, exhiben una molécula mimética del antígeno extraño. Habida cuenta que la inmensa mayoría de las células del cuerpo sintetizan proteínas del CMH I, algunas moléculas del autoantígeno nativo se unen a las del CMH. Estos autoantígenos provocan el ataque de los linfocitos *T* citotóxicos, que ya fueron sensibilizados contra el foráneo antígeno "gemelo". Al propio tiempo, los anticuerpos excitados contra el agente patógeno extraño se unen a las células β y comienzan a atraer macrófagos, proteínas del sistema del complemento y posiblemente células asesinas, o NK.

Hasta este momento, las células β lesionadas podrían resistir el asalto. Mas, entonces, los macrófagos comienzan a presentar el autoantígeno nativo a los linfocitos *T* coadyuvantes (quizá dañando simultáneamente las células con la liberación de la interleucina-1). Llegado este punto, los linfocitos *T* coadyuvantes se multiplican e intensifican la respuesta inmune contra el autoantígeno.

Muy pronto, las células dañadas, tal vez en un esfuerzo por identificarse a sí mismas como propias, avivan la síntesis de moléculas del CMH; no sólo de las moléculas de la clase I, sino también, de acuerdo con recientes observaciones, la síntesis de las de clase II. Las moléculas del CMH I instan el "bombardeo" de los linfocitos *T* citotóxicos. Se cree que, por lo común, la exposición de moléculas del CMH II por células no-inmunitarias provoca la respuesta supresora del sistema inmune. Pero, en este caso, el esfuerzo realizado por la células β fracasa: debido a un defecto genético en el sistema inmune, no se lleva a cabo la respuesta supresora. Antes bien, la exposición de moléculas de la clase II atrae incluso a más linfocitos *T* coadyuvantes y promueve, con más fuerza, la respuesta inmune.

¿Qué sucede entonces? Las células β lesionadas y los macrófagos activados que las rodean producen gran cantidad de radicales libres. Las células β no consiguen ya inhibir su efecto tóxico. Estas podrían también exponer en su superficie, o extruir, proteínas que suelen permanecer secuestradas en el interior celular (como las proteínas de "choque térmico", los gangliósidos citoplasmáticos y la proinsulina). Estas proteínas, al no resultar familiares para el sistema inmune, se las reputa extrañas. Los macrófagos las fagocitan y se las presentan a los linfocitos *T* coadyuvantes, desencadenándose así otra ola

de ataques. Las células β, incapaces de resistir una agresión tan virulenta, no tardan en morir.

Cuantas más células productoras de insulina sucumban, más deberán trabajar las supervivientes para subvenir a la necesidad corporal de insulina. Esta hiperactividad fatiga a las células y quizá las conduce a incrementar la exposición de autoantígenos. En consecuencia, las células β supervivientes mueren también con rapidez acelerada. Al final no quedan suficientes células β para suministrar la insulina que el organismo demanda, y así aparecen "repentinamente" los síntomas de la diabetes.

Para prevenir la destrucción inmunitaria de las células β en los diabéticos potenciales, los médicos han de comenzar por poder identificar los grupos de riesgo. Varios equipos, el nuestro entre ellos, han demostrado que los tests para detectar autoanticuerpos contra el autoantígeno citoplasmático, y contra la insulina, pueden predecir si se desarrollará la enfermedad en parientes cercanos a diabéticos insulino-dependientes (que poseen un alto riesgo de contraer la enfermedad), así como también en grupos de individuos elegidos al azar de la población general.

El autoanticuerpo 64-K constituye otro instrumento de predicción. A la larga, podría incluso convertirse en el método idóneo para un rastreo general. Aparece pronto, y casi todos, si no todos, los individuos que adquieren la diabetes insulino-dependiente lo sintetizan antes de que se manifiesten los síntomas diabéticos. Los tests de laboratorio requeridos para desenmascarar el autoanticuerpo, demasiado costosos y complejos aún, no admiten un uso generalizado. Mientras tanto, nuestro grupo y otros estamos descifrando la secuencia de aminoácidos de la molécula. Una vez obtenida, se podrá diseñar una técnica sencilla de rastreo.

Aunque no existen terapias preventivas aprobadas para el seguimiento de los individuos identificados por medio de tales tests, se están realizando estudios clínicos para determinar posibles tratamientos que permitan abrigar esperanzas. Por ejemplo, en algunos diabéticos de diagnóstico reciente, se protegen las células β supervivientes por medio de agentes inmunosupresores tales como la ciclosporina, los esteroides y la azatioprina, substancias todas que menguan la población de linfocitos *T* o bien inhiben su activación. Alentados por este éxito, varios grupos están considerando la posibilidad de administrar experimentalmente tales drogas como medida preventiva. El planteamiento, sin duda, tendrá éxito, pero los beneficios de la terapia habrán de sopesarse cuidadosamente con sus considerables inconvenientes. Los fármacos, posiblemente, deberán ingerirse durante toda la vida; además, la supresión que producen del sistema inmune incrementará el riesgo de infección y de cáncer.

En el futuro, cuando el autoantígeno desencadenante de la diabetes y su estructura se conozcan con certeza, los investigadores se hallarán capacitados para sintetizar agentes selectivos: que erradiquen específicamente el número moderado de linfocitos *T* que llevan receptores del autoantígeno y, al mismo tiempo, dejen intacto el resto de los linfocitos *T*. En ese contexto, podrían crearse moléculas híbridas, formadas por un autoantígeno y una toxina ligada al mismo. Cuando los linfocitos *T* autorreactivos se unieran al antígeno, la toxina los mataría. Dicho tratamiento, por tanto, no sería perjudicial ya que suprimiría un pequeño subconjunto de linfocitos *T* y no comprometería apenas la capacidad inmunológica del individuo.

Hace unos 10 años, algunos investigadores pensaron en una pronta prevención de la diabetes insulino-dependiente. Hoy existe la creciente convicción de que, antes de que transcurran otros 10 años, la investigación en las causas de la diabetes conducirá a terapias preventivas seguras. Si los efectos destructores de la diabetes insulino-dependiente no pueden anularse con inyecciones de insulina, entonces deberá eliminarse la propia enfermedad.

BIBLIOGRAFIA COMPLEMENTARIA

EPIDEMIOLOGIC APPROACH TO THE ETIOLOGY OF TYPE I DIABETES MELLITUS AND ITS COMPLICATIONS. Andrzej S. Krolewski et al. en *New England Journal of Medicine*, vol. 317, n.º 22, págs. 1390-1398; 26 de noviembre de 1987.

A HYPOTHETICAL MODEL OF THE FOREIGN ANTIGEN BINDING SITE OF CLASS II HISTOCOMPATIBILITY MOLECULES. Jerry H. Brown et al. en *Nature*, vol. 332, n.º 6167, págs. 845-850; 28 de abril de 1988.

HOW, WHEN, AND WHY TO PREDICT IDDM. Noel K. Maclaren en *Diabetes*, vol. 37, págs. 1591-1594; diciembre de 1988.

IMMUNOLOGIC TOLERANCE: COLLABORATING BETWEEN ANTIGEN AND LYMPHOKINES. G.J.V. Nossal en *Science*, vol. 245, n.º 4914, págs. 147-153; 14 de julio de 1989.

GENETIC CONTROL OF AUTOIMMUNITY IN TYPE I DIABETES. John A. Todd en *Immunology Today*, vol. 11, n.º 4, págs. 122-129; abril de 1990.

La Gran Mancha Roja de Júpiter

Longevo e inmenso vórtice de nubes en rotación anticiclónica, mide unas dos veces el diámetro de la Tierra. Su naturaleza y persistencia en la atmósfera de Júpiter siguen representando un desafío para la ciencia

Agustín Sánchez Lavega

Corría el año 1664, algunos años después de que Galileo Galilei apuntara, por vez primera con un telescopio, hacia el planeta Júpiter, cuando el británico Robert Hooke distinguió con su telescopio de 12 pies de longitud focal la presencia de "una pequeña mancha en la mayor de las tres bandas oscuras de Júpiter", tal y como escribía en el primer volumen de las *Philosophical Transactions of the Royal Society*. Un año más tarde, el astrónomo francés, aunque italiano de nacimiento, Giovan Cassini descubría una "mancha permanente" y medía la rotación de Júpiter. Aunque se duda de la naturaleza de la mancha de Hooke, parece que, por su localización, la observada por Cassini y, de forma intermitente durante cuarenta y ocho años, por otros astrónomos de la centuria, correspondió a la conocida hoy por "Gran Mancha Roja" (GRS, de sus iniciales en inglés). Podemos afirmar, pues, que se trata de una formación longeva de la atmósfera de Júpiter, con trescientos años de antigüedad, como mínimo.

Se carece de datos convincentes sobre la presencia de la GRS en el siglo XVIII. Mas, a partir de mediados del siglo XIX, se publican de nuevo dibujos que parecen señalar (a veces de forma indirecta) la existencia de la Mancha. Desde finales del siglo pasado poseemos datos concretos y continuos del movimiento y tamaño de tan peculiar detalle.

Basta en general con un pequeño telescopio y con unas medianas condiciones de estabilidad en nuestra atmósfera para observar la presencia de ese fenomeno sobre el planeta. Suele aparecer como un óvalo, ligeramente rojizo, incrustado entre una franja blanca —Zona Tropical Sur— y una franja oscura —Banda Ecuatorial Sur—, cerca del paralelo 23 grados del hemisferio Sur de Júpiter. Se nos ofrecerá más o menos prominente según sea el estado de actividad de esas dos regiones. Cuando la Banda Ecuatorial Sur (SEB) apenas existe y la Zona Tropical Sur (STRZ) es blanca y uniforme, destaca la presencia de la Mancha, como sucediera durante la primera mitad de la década de los años setenta (y desde el verano de 1989). Con este aspecto se presentó ante las cámaras de los vehículos espaciales *Pioneer 10* y *11* en los años 1973 y 1974. Por contra, apenas se percibe cuando la SEB y parte de la STRZ se vuelven oscuras y turbulentas; la Mancha constituye, entonces, una oquedad entre ambas. Tal es lo que ha venido sucediendo desde mediados del año 1975 hasta mediados de 1989; así se mostró a los equipos embarcados en las naves *Voyager 1* y *2* en 1979.

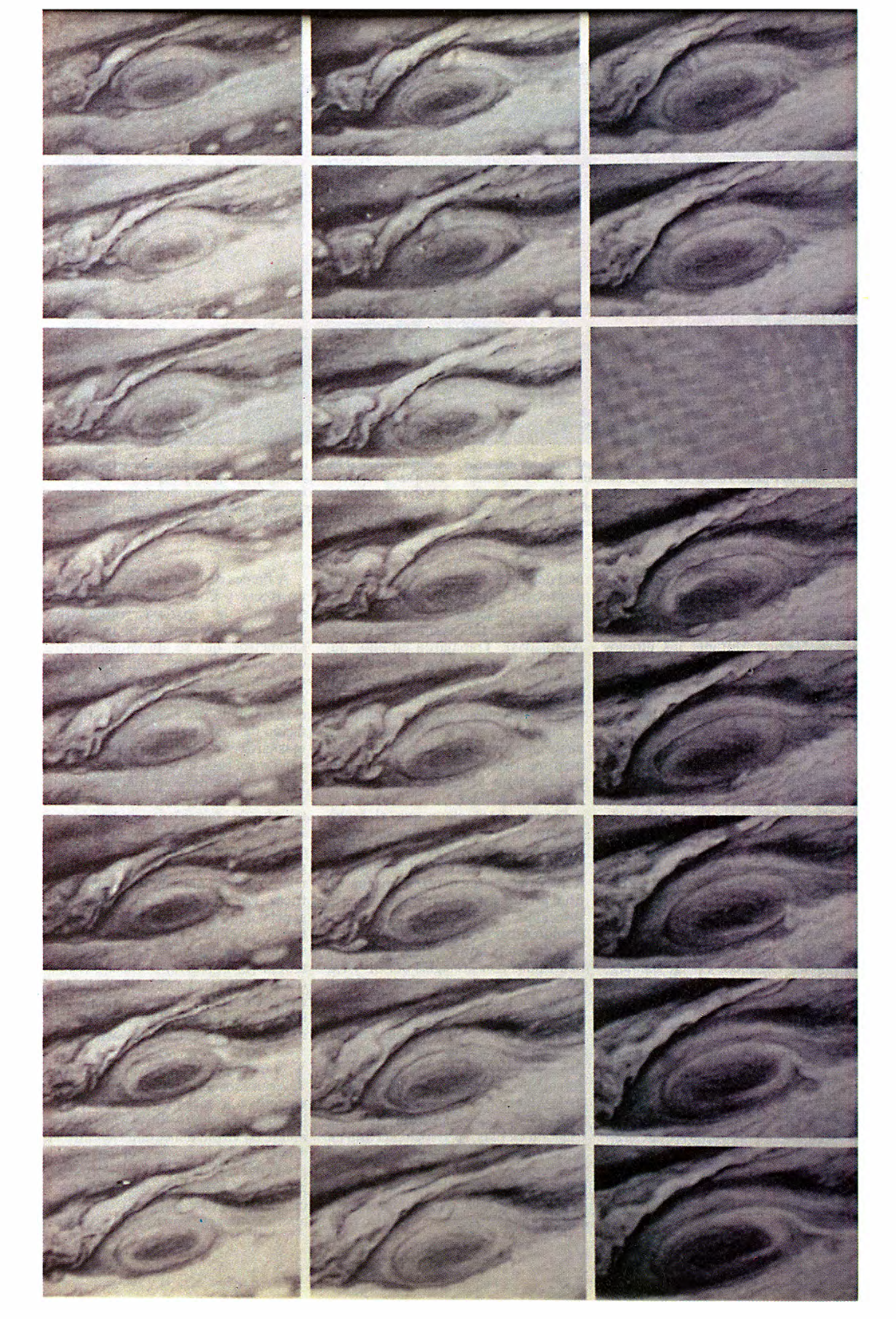
Conviene, antes de seguir adelante, recordar algunas de las principales características del planeta Júpiter. Es el de mayor volumen del sistema solar, posee un radio ecuatorial de 71.398 kilómetros y consta primordialmente (lo mismo que los otros planetas gigantes, Saturno, Urano y Neptuno) de hidrógeno y helio. Dotado de una densidad media de 1,5 gramos por centímetro cúbico, cuanto vemos sobre su "superficie" son formaciones nubosas de variado colorido y reflectividad, situadas sobre una atmósfera profunda. Las nubes se desplazan a lo largo de los paralelos en corrientes hacia el Este y Oeste, alternantes con la latitud.

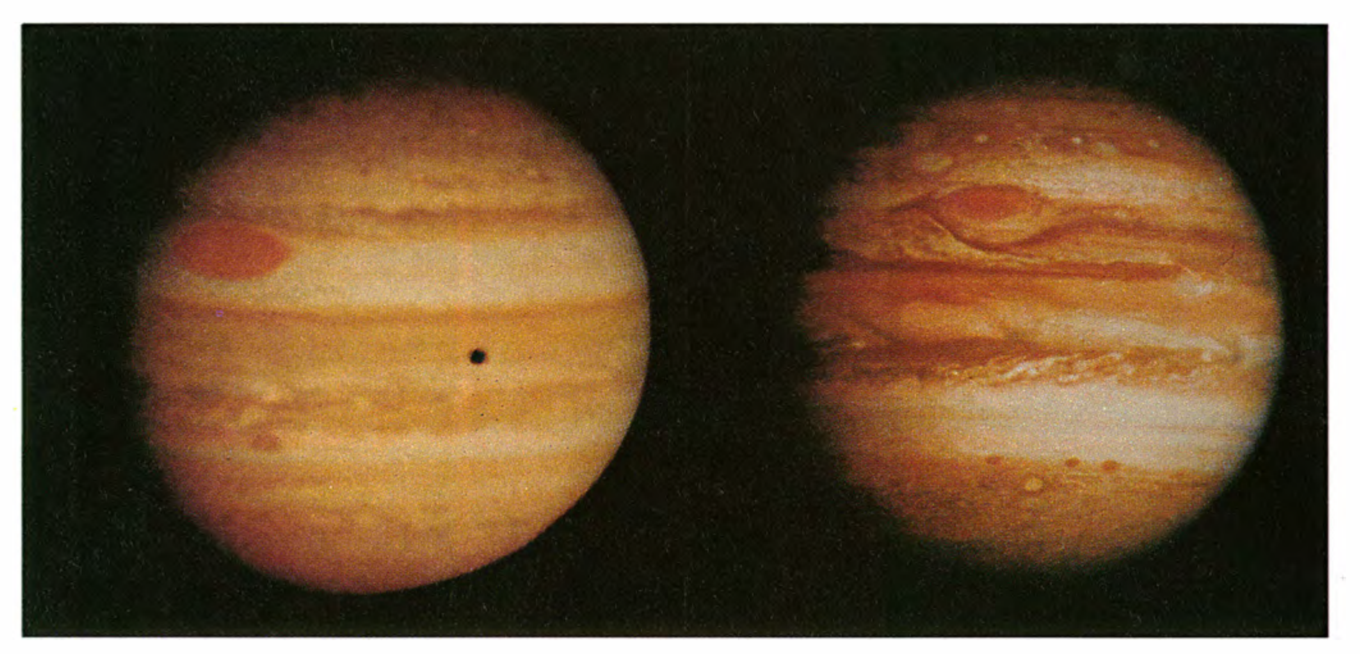
De las variaciones en la emisión radioeléctrica del planeta inferimos, además, que el interior de Júpiter da una vuelta sobre sí mismo en unas 9 horas 55 minutos 30 segundos, lo que le convierte en el planeta que gira más rápido sobre su eje de todo el sistema solar. Dada su gran distancia al Sol, la cantidad de calor que recibe apenas alcanza un cuatro por ciento de la que se recibe en la Tierra. A esta fuente de energía, hay que añadir el calor que Júpiter genera en su interior, cifrado en unas 1,7 veces el recibido del Sol. En estas condiciones, la temperatura en las nubes superiores (donde la presión atmosférica es del orden de 1 bar) se cifra en unos 120 grados Celsius bajo cero. Su gran tamaño, la alta velocidad angular de rotación y el calor disponible en la atmósfera son probablemente los parámetros que condicionan la dinámica atmosférica de Júpiter y, por consiguiente, la presencia de gigantescos vórtices como la GRS [véase "Júpiter y Saturno", por A. P. Ingersoll; INVESTIGACIÓN Y CIENCIA, febrero de 1982].

De las mediciones del tamaño de la GRS realizadas en el último siglo se desprende que, aun manteniendo su forma elíptica, sus dimensiones no son constantes. Los dibujos realizados en 1879 por diversos observadores nos permiten deducir que el eje menor (en dirección Norte-Sur) medía en aquella época 13.000 kilómetros, mientras que el mayor (en la dirección Este-Oeste) era de 40.000 kilómetros. De entonces acá se ha contraído fundamentalmente en la direc-

AGUSTIN SANCHEZ LAVEGA es doctor en ciencias físicas por la Universidad del País Vasco, de cuya Escuela Superior de Ingenieros es profesor. Especialista en atmósferas planetarias, trabajó de 1980 a 1987 en el Centro Astronómico Hispano Alemán (Observatorio de Calar Alto) en Almería. Colaboró en la investigación de las atmósferas de Júpiter y Saturno con el Instituto de Astrofísica de Andalucía en Granada y con el Observatorio de París-Meudon.

1. ROTACION ANTICICLONICA de la Mancha Roja e interacción con otros óvalos menores que se aproximan por el Este. La secuencia obtenida, a lo largo de varias rotaciones del planeta, en 1979 muestra cómo los pequeños vórtices blancos circulan por la periferia de la Mancha para acabar engullidos por ella.





2. JUPITER Y SU GRAN MANCHA ROJA, según imágenes obtenidas por el vehículo espacial *Pioneer 10* el dos de diciembre de 1973 (*izquierda*) y por el *Voyager 1* el primero de febrero de 1979 (*derecha*). Las grandes diferencias en la apariencia visible de las bandas reflejan el cambio que se produjo en el planeta en el verano de 1975. La Mancha Roja destaca por contraste frente al fondo blanco: un óvalo aislado cuando la Banda Ecuatorial Sur desaparece (*izquierda*); queda enmascarada cuando dicha banda y la Zona Tropical Sur se encuentran en estado activo, lo que origina complejas estructuras nubosas oscuras (*derecha*). La Mancha mide actualmente unos 22.000 kilómetros (Este-Oeste) por 11.000 kilómetros (Norte-Sur).

ción zonal, es decir, a lo largo de su eje mayor; hoy mide unos 25.000 kilómetros. Segun R. Beebe y L. A. Youngblood, de la Universidad de Nuevo México, durante los últimos cincuenta años la Mancha se ha empequeñecido a un ritmo anual de unos 100 kilómetros, aunque en los últimos años esta contracción parece haberse ralentizado e incluso detenido, según hemos podido comprobar en nuestras propias observaciones realizadas en el Observatorio de Calar Alto, Almería.

La Gran Mancha Roja no es una estructura estática de la atmósfera de Júpiter. Situada entre dos corrientes de vientos zonales, una por el Norte, a 216 kilómetros por hora hacia el Oeste, y otra por el Sur, a 144 kilómetros por hora hacia el Este, la Mancha se mueve hacia el Oeste con una velocidad media de 11 kilómetros por hora, como si fuese rodando entre ambas. En 1968, H. G. Solberg, de la Universidad de Nuevo México, descubrió que la GRS presentaba un movimiento oscilatorio, superpuesto a su movimiento medio, de un período de unos 90 días y una amplitud de 1000 kilómetros.

Sin embargo, los registros de su posición recopilados durante más de cien años, y publicados por B. M. Peek en *El planeta Júpiter*, muestran profundas irregularidades en el movimiento. En las últimas tres décadas se han llevado a cabo investigaciones más cuidadosas de los cambios de velocidad de la GRS, apoyadas en el análisis de fotografías telescópicas de alta resolución y, más recientemente, de imágenes tomadas con cámaras electrónicas CCD (Dispositivos de Cargas Acopladas). Trabajaron en ello, durante la década de los sesenta y setenta, E. J. Reese y R. Beebe, entre otros, de la Universidad de Nuevo México. En nuestro propio estudio de la cinemática de la Mancha hemos analizado imágenes tomadas por numerosos observadores de diferentes países en un programa de cooperación internacional.

Parece ahora claro que el movimiento de la GRS sufre una alteración transitoria cuando ésta entra en interacción con otras estructuras atmosféricas que se encuentren en una latitud próxima y que cursen a diferente velocidad. Así sucede, por ejemplo, cuando otros óvalos blancos menores situados en la latitud 33 grados Sur, y bautizados con las siglas BC, DE y FA, adelantan a la GRS en su movimiento hacia el Este a lo largo de su flanco Sur. El desplazamiento hacia el Oeste de la GRS sufre cambios importantes cuando ésta entra en contacto con un sistema complejo nuboso que de vez en cuando evoluciona en las latitudes de la Mancha Roja: con la Perturbación de la Zona Tropical Sur. En cuestión de días, su velocidad varía en unos cuantos kilómetros por hora, acelerándose en un comienzo y decelerándose luego al concluir la interacción, según hemos podido comprobar en numerosas ocasiones.

A finales de la década de los sesenta, Reese y B. A. Smith, de la Universidad de Arizona, observaron cómo manchas oscuras, de unos 6000 kilómetros, que se aproximaban a la GRS por sus bordes, eran absorbidas por ésta, tras haber circulado a lo largo de su periferia; ello indicaba la existencia de un movimiento de rotación de la GRS sobre sí misma. Años más tarde, en 1979, el equipo de imagen de los vehículos *Voyager 1* y *2* dirigido por Smith confirmaba la rotación de la Mancha. Según el mapa detallado de velocidades obtenido en ese proyecto, la Mancha gira anticiclónicamente alrededor de su centro geométrico, siendo las líneas de corriente elipses perfectamente concéntricas. Por su parte, la velocidad tangencial depende de la distancia al centro de la elipse, pero apenas se observan movimientos radiales hacia afuera (divergentes) o hacia adentro (convergentes). La velocidad tangencial alcanza los 430 kilómetros por hora en sus extremos Este y Oeste, y unos 100 en los extremos Norte y Sur.

La explicación dinámica que merecen esos datos numéricos es la siguiente: a lo largo del eje menor de la elipse, el viento fluye geostróficamente en virtud del equilibrio entre la fuerza de Coriolis y la ejercida por el gradiente horizontal de la presión; por contra, a lo largo del eje mayor, el viento se dice de gradiente, pues ahora la fuerza centrífuga, importan-

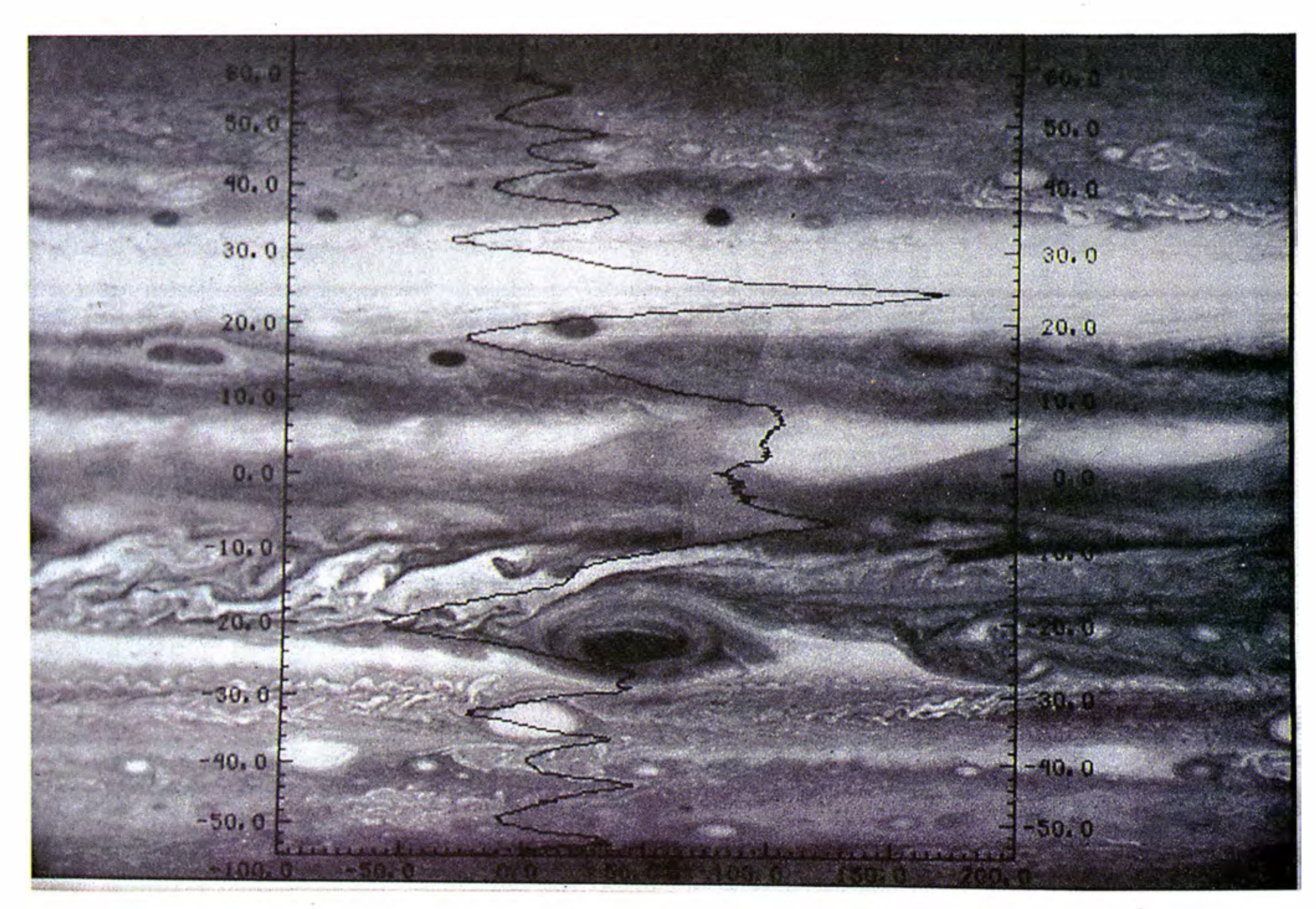
te, interviene en el equilibrio de fuerzas.

No todas las formaciones nubosas que se acercan a la Mancha Roja por los paralelos cercanos son absorbidos por ella. En muchos casos parece justamente lo contrario; actúa a modo de barrera infranqueable, desviándolos de su latitud. Hemos observado este comportamiento en algunos sistemas nubosos (casi siempre pequeños vórtices) que avanzan desde el Este hacia la Mancha; se desvían a lo largo del flanco Norte en la región de la "bahía de la Mancha Roja". En otras ocasiones, los dirige hacia latitudes más sureñas; allí, arrastrados por las corrientes hacia el Este, dan origen a un flujo de recirculación. Ese cuadro es el que se observa, por ejemplo, en la región de la Perturbación Tropical Sur.

En 1975 se produjo un profundo cambio en la textura de las bandas del planeta. Observamos que la Mancha Roja constituía la causa directa de la formación de un sistema anticiclónico de larga vida, situado a sólo seis grados del Ecuador. Al Oeste, sobre la Banda Ecuatorial Sur, emergían nubes blancas brillantes, probablemente de origen convectivo, que, impulsadas por una fuerte corriente del Oeste, se movían hacia la GRS. Al alcanzarla, se comprimían por el borde Norte de la Mancha y desviaban hacia latitudes ecuatoriales. Según un modelo que sugerí años más tarde, el sistema nuboso ecuatorial debió adquirir vorticidad anticiclónica al cambiar de latitud. Moviéndose muy rápido hacia el Este, a unos 325 kilómetros por hora, el anticiclón ecuatorial se encontraba con la GRS cada cincuenta días, alimentándose de las nuevas nubes blancas que, en continua formación, eran expelidas hacia el Norte por la circulación de la Mancha. Mediante este mecanismo, la estructura, de unos 10.000 kilómetros de tamaño, se ha mantenido visible durante los últimos trece años.

La descripción morfológica y los aspectos cinemáticos y dinámicos que hemos comentado nos dibujan una imagen bidimensional (horizontal) de la Mancha Roja correspondiente a una presión atmosférica de medio bar. Pero, ¿qué acontece por encima y por debajo del nivel de la capa superior de nubes? Nuestro conocimiento de la estructura vertical proviene fundamentalmente de cuatro tipos de estudios indirectos: en primer lugar, la reflectividad en la región de longitudes de onda desde el ultravioleta hasta el infrarrojo próximo, incluido tanto el continuo (dominado por la absorción por partículas) como la espectroscopía de las bandas y líneas de absorción de diferentes compuestos (sobre todo los gases metano, amoníaco e hidrógeno); en segundo lugar, la polarización de la luz reflejada en función de la longitud de onda y del ángulo de observación; en tercer lugar, emisión térmica (espectroscopía e imágenes) en el infrarrojo y en ondas de radio, y, por último, la extinción de las ondas de radio enviadas por las sondas Voyager, a medida que éstas quedaban ocultadas por el borde del planeta.

Comenzamos por estas dos últimas. El análisis de la emisión infrarroja y de la radio-ocultación permite efectuar un sondeo vertical del campo tér-



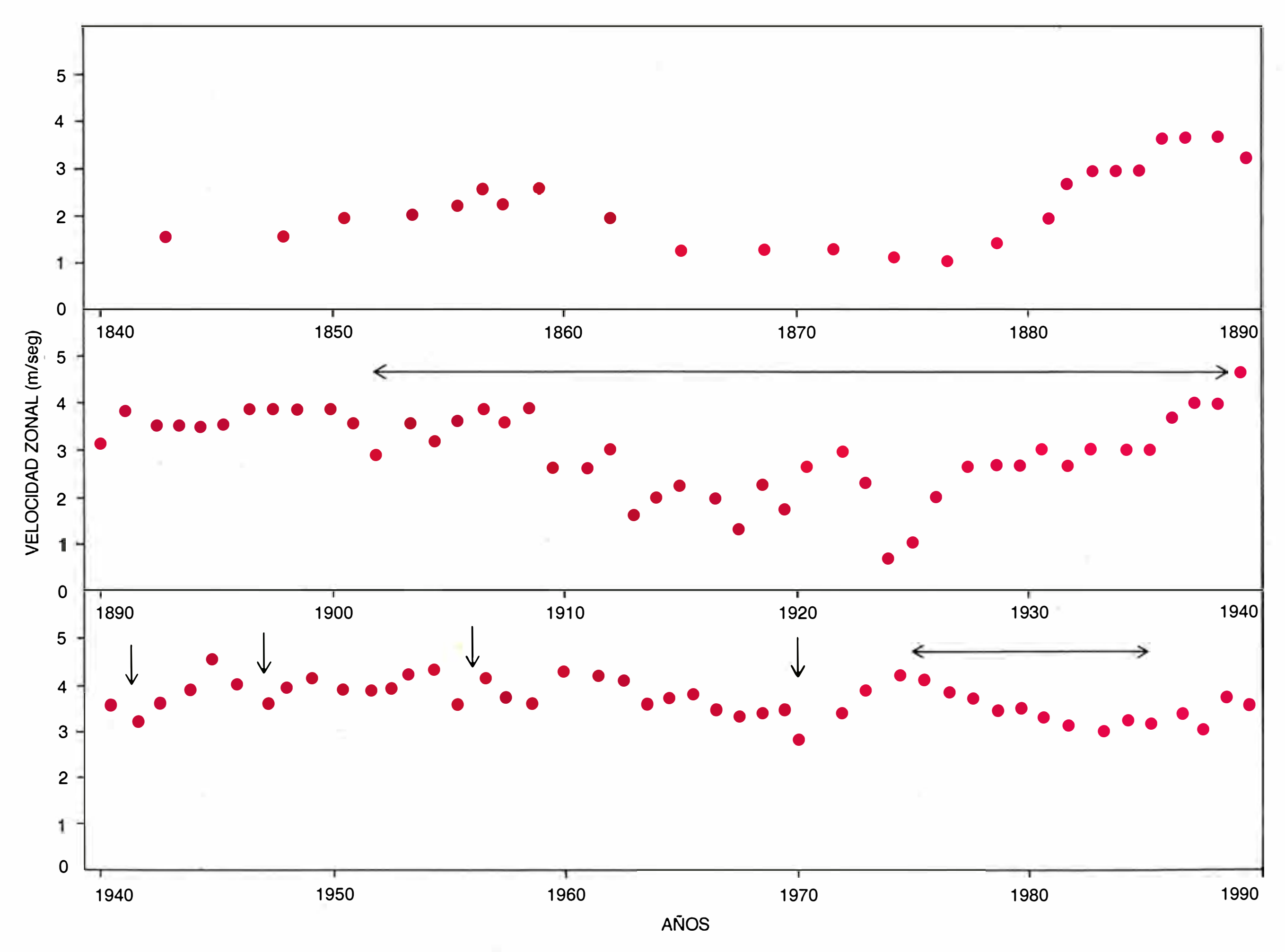
3. LATITUD Y VIENTOS ZONALES en Júpiter. Medidos éstos en metros por segundo, se han determinado a partir de las imágenes proporcionadas por *Voyager 1* y *2*. La Mancha Roja se encuentra limitada al Norte por una corriente hacia el Oeste de unos 60 metros por segundo (216 kilómetros por hora) y al Sur por otra hacia el Este de unos 40 metros por segundo (144 kilómetros por hora). Al Oeste de la Mancha (*izquierda de la imagen*), se aprecia la región turbulenta de nubes blancas que, en su desplazamiento hacia el Este, son comprimidas por la circulación del vórtice, realimentando el anticiclón ecuatorial observado en los últimos años. Al Este de la GRS (*derecha*) se encuentra una región oscura y curvada: la Perturbación Tropical Sur, que se ha detectado a veces en la región de la Mancha. Su mutua interacción produce alteraciones en sus respectivos movimientos.

mico, extrayendo las temperaturas aproximadamente entre los niveles de presión de 1 milibar y unos 2 a 4 bares (lo que equivale a un rango de unos 200 kilómetros de altura). En promedio, cerca de los 140 milibares se encuentra la tropopausa de Júpiter, es decir, el nivel de temperatura atmosférico más bajo (unos 163 grados Celsius bajo cero). Para presiones por debajo de este valor (a medida que ascendemos), la temperatura crece; lo mismo sucede para presiones mayores (a medida que descendemos). Pues bien, F. M. Flasar y B. Conrath, del Centro de Vuelos Espaciales Goddard de la NASA, han podido examinar con precisión, gracias a los datos suministrados por los vehículos *Voyager 1* y *2*, la estructura térmica por encima del techo nuboso de la Mancha Roja. Esta se caracteriza por poseer una tropopausa relativamente fría con respecto a sus alrededores, mientras que su estratosfera se encuentra bastante caliente.

Para estudiar qué sucede por debajo de las nubes superiores, conviene acudir a la radiación de cinco micrometros de longitud de onda, en donde existe una "ventana" a través de la cual se escapa la radiación desde niveles profundos (unos cinco bares de presión). La atmósfera es transparente (ningún gas allí presente absorbe dicha radiación) y sólo las nubes muestran opacidad entre los dos y cuatro bares. Las imágenes telescópicas de la GRS obtenidas en cinco micrometros por R. J. Terrile, del Laboratorio de Propulsión a Chorro en Pasadena (JPL), la muestran como un objeto frío, señal de que se halla cubierta por espesas nubes, salvo en el borde exterior de la elipse; en éste se aprecia un anillo caliente de alta emisión infrarroja que corresponde a una región de escasa nubosidad.

Para determinar la estructura vertical nubosa podemos acudir a las imágenes tomadas en las anchas bandas de absorción del metano en el rojo e infrarrojo próximo, mediante el empleo de filtros interferenciales. R. West, del JPL, y M. G. Tomasko, de la Universidad de Arizona, han unido esta técnica al análisis espectroscópico de las delicadas líneas de absorción del hidrógeno, y han llegado a la conclusión de que la Mancha es una estructura alta de la atmósfera. El techo de nubes se encuentra en el nivel 0,18 bar frente al nivel 0,34 bar para las regiones adyacentes, es decir, unos 13 kilómetros por encima de ellas. Los registros de la polarización de la luz reflejada avalan esta conclusión. Sin embargo, el perímetro de la Mancha aparece como una región desprovista de nubes superiores, en sintonía con lo deducido del análisis infrarrojo. De todos estos datos, podemos concluir que la GRS tiene, como mínimo, una extensión vertical de 150 kilómetros, al menos hasta donde podemos sondear con técnicas indirectas.

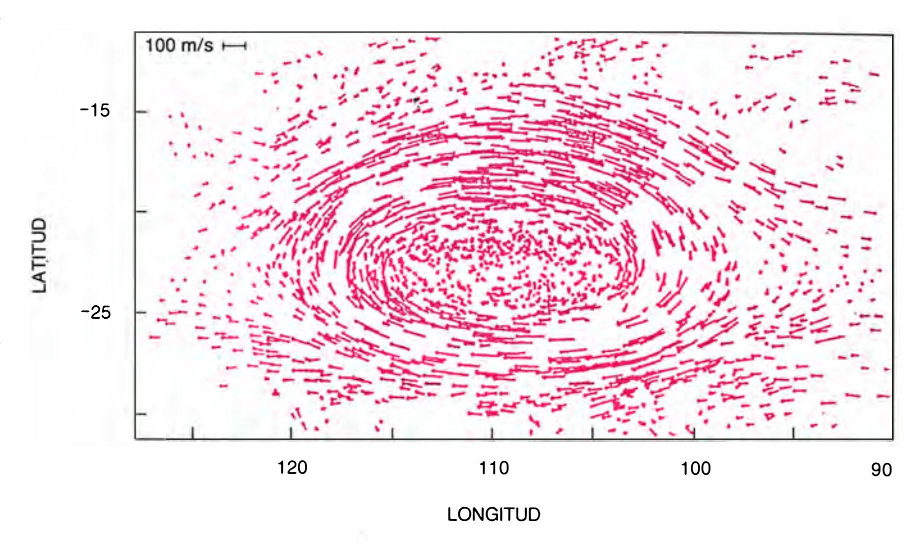
Antes de pasar a describir los modelos dinámicos propuestos, nos queda aún por hablar de la característica

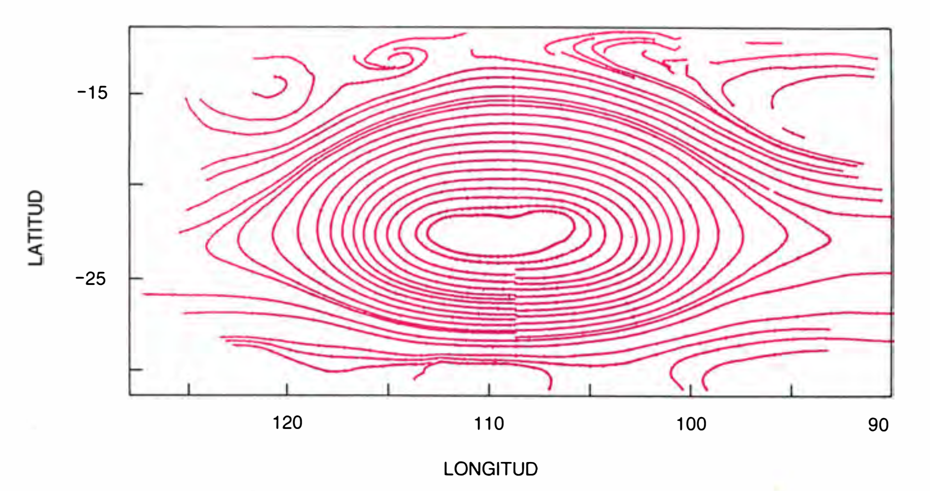


4. VELOCIDAD ZONAL DE LA MANCHA ROJA entre 1840 y 1990. El gráfico muestra medias anuales durante 130 años de la velocidad hacia el Oeste de la GRS. Los mínimos corresponden aproximadamente a las épocas durante las cuales la Perturbación Tropical Sur estuvo activa (señalado con flechas horizontales y verticales), alterando el monótono movimiento de la Mancha. Esta sufre, además, cambios bruscos de corta duración en su desplazamiento, cuando colisiona con otros detalles próximos. Los datos de los últimos quince años corresponden a medidas efectuadas por el autor.

que da lugar a su nombre: su color rojo. Curiosamente, la inspección telescópica visual puede provocar más de una decepción, ya que en general apenas se vislumbra el tinte rojizo. Este adquiere mayor nitidez durante las épocas de desaparición de la Banda Ecuatorial Sur, cuando la Mancha resalta en su singularidad. El uso de filtros violetas y azules destacan su contraste frente a las regiones próximas, todo lo contrario de lo que sucede si colocamos un filtro rojo. Hay que buscar el origen de la tonalidad rojiza en algún colorante, formado por procesos de desequilibrio químico, que se mezcle con las nubes superiores blancas de amoníaco. Su naturaleza sigue siendo un misterio, pues los diferentes agentes propuestos escapan a la detección. En general, el origen del colorido de las nubes de los planetas gigantes constituye uno de los mayores retos de los investigadores planetarios.

Una teoría coherente de la Mancha Roja debería explicar todas las propiedades observadas y descritas antes. Sin embargo, los modelos propuestos hasta la fecha presentan importantes lagunas por nuestro desconocimiento de aspectos fundamentales de la atmósfera del planeta. En este sentido, se carece de datos fidedignos sobre la estructura térmica local de la Mancha, por debajo de la capa superior de nubes, o de la naturaleza de las capas nubosas y del papel que desempeñan los posibles efectos de liberación de calor latente. O lo que es más importante: no sabemos quién controla globalmente la dinámica atmosférica en la región de las nubes. ¿Lo hace la fuente interna de energía? ¿Depende de la radiación solar o de ambas? Describiremos, en lo que sigue, los aciertos y las dudas en que se mueven las teorías enunciadas sobre la GRS.



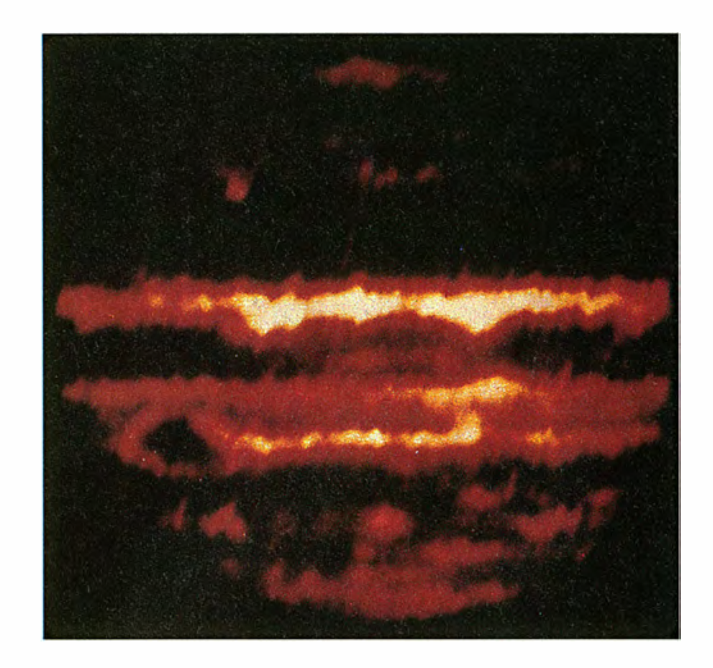


5. MAPA HORIZONTAL DE VELOCIDADES Y TRAYECTORIAS (*arriba y abajo respectivamente*), **obtenidas a partir de las imágenes suministradas por los Voyager. En el diagrama superior cada punto señala la posición en longitud y latitud del detalle nuboso medido; su dirección de movimiento viene dada por el vector y la intensidad del viento por su longitud (escala en metros por segundo). Nótese en el diagrama inferior cómo la circulación de la GRS tiene lugar en forma de elipses concéntricas alrededor del centro geométrico. Mientras que la velocidad tangencial alcanza en los extremos Este y Oeste los 430 kilómetros por hora y unos 100 kilómetros por hora en los bordes Norte y Sur, la velocidad radial es prácticamente nula. Esto parece indicar que, en el nivel de las nubes, apenas existen movimientos divergentes o convergentes de gran escala en el interior de la Mancha.**

Hasta comienzos de la década de los sesenta, la interpretación más común presentaba la Mancha Roja como un objeto flotante. G. W. Hough avanzó esa idea en 1881, R. Wildt la desarrollaría en 1939. En esa línea de razonamiento, la Mancha era un cuerpo sólido o líquido de baja densidad que vagaba por la atmósfera de Júpiter.

Pero ya en 1961, R. Hide, entonces en el King's College de Londres, sugirió que la Mancha podría corresponder a la parte superior de una "columna de Taylor". Estas columnas se forman cuando un fluido en rápida rotación supera un obstáculo, no muy profundo, que se mueve lentamente en dirección perpendicular al fluido. Si las velocidades relativas son apropiadas, se origina verticalmente una columna de estancamiento sobre el objeto. El movimiento y la estructura que se forma en la parte superior de la columna se aproximan al observado en la GRS. En particular, se reproduce la estructura de "bahía", característica del flanco Norte de la Mancha. Aunque de fácil creación en el laboratorio, no parece plausible la formación de esas columnas en condiciones geofísicas (que incluyan, por ejemplo, estratificación atmosférica). En cualquier caso, al margen de las objeciones cinemáticas, dinámicas y del campo térmico que se conocen, ¿cuál sería el obstáculo subyacente a las nubes superiores capaz de producir la estructura en columna? Por otro lado, no se ha observado ningún movimiento generalizado en latitud de la Mancha Roja equivalente al movimiento en ángulo recto que sufre el objeto de la simulación recreada en el laboratorio.

Uno de los pioneros de la moderna investigación planetaria, ya fallecido, G. P. Kuiper, de la Universidad de Arizona, sugirió, al comienzo de los setenta, que la responsable de la GRS podría hallarse quizás en la convección húmeda en pequeña escala. El vórtice vendría a ser aquí un fenómeno parecido a los huracanes tropicales terrestres. La convergencia de masas de gas húmedas de vapor de agua en niveles profundos cercanos a los 7 bar y la subsiguiente liberación de calor latente constituirían su fuen-





te de energía. La hipótesis del huracán, en boga durante esa década, recibió años más tarde sólido respaldo de F. M. Flasar, B. J. Conrath y P. Gierasch, que se apoyaron en las mediciones térmicas efectuadas con los Voyager. De acuerdo con las mismas, al nivel de las nubes, el núcleo de la GRS sería un centro de altas presiones y, por debajo, habría una región caliente. Añaden, además, que las bajas temperaturas estratosféricas detectadas podrían estar asociadas a un vigoroso movimiento de ascensión.

No obstante las semejanzas entre la estructura térmica vertical de un huracán y la deducida para la GRS, la dirección de la circulación en la Mancha, anticiclónica, y concentrada en la periferia, parecen invalidar la hipótesis del huracán. Además, resulta difícil justificar la convergencia de masas de gas húmedo en los niveles inferiores a la capa de nubes, ya que se carece de un suelo o superficie sólida capaz de catalizarlo.

John Scott Russell, ingeniero naval inglés, había observado en 1834 la formación y desplazamiento de una

6. DIVERSOS ASPECTOS DE JUPITER y su Mancha Roja, cuando se observa en diferentes longitudes de onda. En la parte superior pueden compararse la imagen obtenida desde el Observatorio de Monte Palomar en el infrarrojo en cinco micrometros, correspondiente a la emisión térmica en niveles por debajo de la capa visible de nubes (*izquierda*), con la obtenida en luz visible con el *Voyager 1*, ambas en el año 1979 (*derecha*). En la imagen infrarroja las regiones calientes aparecen brillantes, mientras que las frías son las oscuras. El centro de la GRS aparece como una región fría mientras que, por su periferia, se observa un anillo caliente. En la parte inferior pueden compararse dos imágenes obtenidas en el Observatorio del Pic du Midi en 1987 con un filtro que aísla la banda de absorción por el gas metano en 8900 A (*derecha*) con la obtenida con un filtro en el continuo adyacente (*izquierda*). El hecho de que la Mancha Roja destaque por su brillo (alta reflectividad) en la profunda banda del metano sugiere que su techo de nubes se encuentra alto en la atmósfera, unos 15 kilómetros por encima con respecto a sus alrededores.

onda peculiar, de una única cresta, que, sola y sin cambiar de forma, descendía canal abajo hasta el mar. A finales del siglo XIX, D. J. Korteweg y G. De Vries idearon una ecuación matemática que describía el comportamiento de esta onda, ecuación denominada en su honor *KdV*. Se ha comprobado ya la existencia de este tipo de ondas en acústica, óptica, oceanografía y otras. Se las conoce por ondas solitarias o "solitones". En 1976, T. Maxworthy y L. C. Redekopp, de la Universidad de California en Los Angeles, propusieron que la GRS podía ser la manifestación visible en la capa de nubes de un solitón.

En los planetas gigantes, y en particular en Júpiter, estas ondas podrían formarse si concurriera una fuerte estratificación en la densidad atmosférica. Situadas entre dos corrientes de vientos que cursan en direcciones contrarias, las ondas se originarían al curvarse y cerrarse las líneas de corriente en sus bordes Este y Oeste; crearían así estructuras ovaladas en una banda de latitud de alto gradiente meridional en lo concerniente a la velocidad zonal. Ahora bien, cuando se produce una interacción entre solitones, éstos no cambian ni su forma ni su amplitud; se asiste sólo a un corrimiento de fase que semeja la aceleración de uno de ellos. Curiosamente este tipo de interacción es el observado entre la GRS y la Perturbación Tropical Sur. Además, el mapa de líneas de corriente observado en sus alrededores concuerda con el que predice la teoría.

Mas, a pesar de estos parecidos, encontramos graves limitaciones del modelo. Habría que explicar la persistencia de la onda solitaria frente a la acción de los efectos disipativos, tan importantes en la turbulenta atmósfera de Júpiter, y justificar también la estructura térmica vertical observada.

En este sentido, diez años después S. V. Antipov y sus colaboradores, del Instituto Kurchatov de la Energía Atómica en Moscú, reprodujeron en el laboratorio, mediante un original experimento que simulaba la estructura horizontal de los vientos en Júpiter, la presencia de un vórtice solitón aislado. A diferencia del modelo propuesto por Maxworthy y Redekopp, en el modelo de laboratorio del equipo soviético el vórtice solitario interaccionaba y se mezclaba con otros vórtices similares que se formaban entre las corrientes que fluían en dirección opuesta, y mantenía su naturaleza solitaria sin desintegrarse por los efectos destructivos de la fricción.

En 1979, G. P. Williams, del Laboratorio de Dinámica de Fluidos Geofísicos de la Universidad de Princeton, realizó extensos cálculos numéricos en el ordenador para resolver las ecuaciones hidrodinámicas que representaran los procesos físicos de la atmósfera de Júpiter. Partía de la hipótesis según la cual dichos procesos eran del mismo tenor que los que operan en nuestra propia atmósfera, salvadas las diferencias de parámetros que importan más en uno u otro planeta. Se supone, por ejemplo, que el caldeo diferencial solar (variable a lo largo del año de Júpiter en cada región del planeta) controla los movimientos atmosféricos. Los resultados de Williams reproducían los sistemas alternantes de vientos zonales y el esquema característico de bandas y zonas. Es notoria, además, la presencia de óvalos situados entre las corrientes, alguno de ellos aislado; por su aspecto (forma, tamaño y estructura horizontal de las líneas de corriente en su entorno), esa recreación ofrece un estrecho parecido con la GRS. Los vórtices se forman en el modelo como resultado de la inestabilidad baroclínica, extrayendo energía de las diferencias de temperatura entre ecuador y polo. Es conocida la intervención de este mecanismo en la





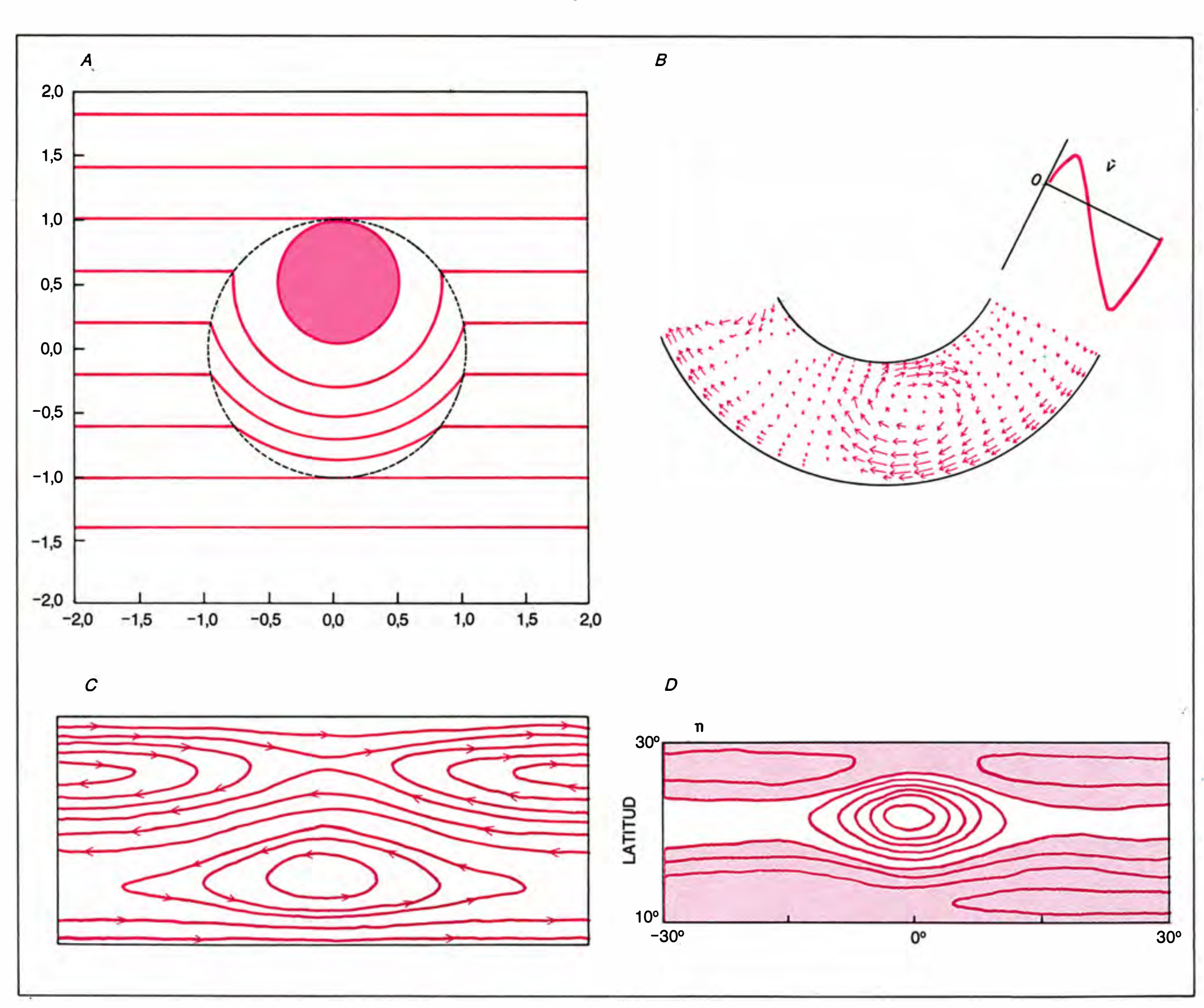
7. DOS EPOCAS DE DIFERENTE ACTIVIDAD de la Banda Ecuatorial Sur (SEB) y de la Zona Tropical Sur (STrZ). A la izquierda, la imagen obtenida por el vehículo espacial *Pioneer 11* el tres de diciembre de 1974 durante una época de calma; a la derecha, imagen del *Voyager 2* el siete de julio de 1979 durante una época de actividad. En ambos casos, por el flanco Sur de la GRS se aprecia el óvalo blanco conocido por las siglas DE situado en la latitud 33 grados Sur. Otros dos óvalos denominados BC y FA acompañan al anterior (no visibles en las imágenes). Los tres se desplazan hacia el Este, encontrándose con la Mancha periódicamente, lo que produce una leve alteración de su movimiento. Se trata de vórtices anticiclónicos muy semejantes en sus propiedades dinámicas a la Gran Mancha Roja de Júpiter. Tan sólo les separa el color, algo aún sin explicación.

generación de las borrascas y anticiclones de las latitudes medias de la atmósfera terrestre; podría actuar también en la de Júpiter.

En colaboraciones posteriores con T. Yamagata y R. Wilson, Williams realizó nuevas y más elaboradas simulaciones numéricas. Según la estratificación atmosférica y la latitud, pueden, muestran esos estudios, emerger vórtices robustos y aislados a partir de inestabilidades en la cizalla del viento zonal. Las direcciones opuestas de las corrientes ayudan a confinar el vórtice. Su circulación es anticiclónica y su estabilidad alta. Además, cuando interacciona con otros vórtices menores se produce su mezcla. Se ha criticado a este tipo de modelos la omisión de la fuente de energía interna del planeta. Según ellos, sobre la capa nubosa superior en donde se originan la GRS y las demás estructuras meteorológicas que observamos, tan sólo influye el calor recibido en forma de radiación solar. Habría que explicar por qué no lo hace el calor proveniente del interior, máxime teniendo en cuenta el vigor de esa fuente de energía. La hipótesis de Williams y sus colegas supone también que la frontera inferior de esa "capa del tiempo" es rígida e impermeable, algo de nuevo difícil de sostener debido a la presencia del transporte turbulento por convección del calor interno.

El papel del proceso baroclínico (es decir, el gradiente meridional en la temperatura) en la formación de vórtices en las atmósferas de los planetas gigantes ha sido estudiado desde otra perspectiva por P. L. Read y R. Hide, del Laboratorio de Dinámica de Fluidos Geofísicos perteneciente a la Oficina Meteorológica Británica. Su análisis parte de ensayos de laboratorio, en los que se investiga el comportamiento de un fluido en un depósito cilíndrico en rotación y sometido a un fuerte gradiente térmico entre sus paredes. En el experimento se producen remolinos aislados, regulares y estables. Según sea el signo del gradiente, los remolinos giran ciclónicamente o anticiclónicamente. De esta forma, la estructura horizontal evoca, una vez más, la de la Mancha en su mapa de velocidades y líneas de corriente.

A. P. Ingersoll y P. G. Cuong, del Instituto de Tecnología de California, propusieron en 1981 que la Mancha Roja era un "modon", término empleado por los oceanógrafos para designar los anillos de circulación cerrada de agua que se originan en las corrientes marinas, en particular los desarrollados en la corriente



8. MODELOS DE LA MANCHA ROJA. Se presentan en diagramas horizontales de latitud frente a longitud: (*a*) líneas de corriente en la parte superior de una columna de Taylor, originada cuando un fluido en rápida rotación sobrepasa un obstáculo no muy profundo que se mueve perpendicularmente respecto del fluido; (*b*) campo de velocidades correspondiente a un vórtice baroclínico observado en el laboratorio en un depósito cilíndrico en alta rotación cuando sus paredes se encuentan a diferente temperatura; (*c*) líneas de corriente en un solitón, u onda solitaria de una sola cresta que se desplaza, sin interaccionar, en un fluido con cizalla en su velocidad; (*d*) vórtice anticiclónico generado en el ordenador al integrar ecuaciones dinámicas semejantes a las usadas para simular la circulación atmosférica de nuestro planeta. El vórtice, coherente y estable, crece.

del Golfo. Los modones podrían formarse en condiciones normales de estratificación vertical y suponiendo que las corrientes de vientos zonales alternantes se extendiesen en profundidad por la atmósfera. Nacerían y crecerían, sin embargo, a modo de perturbación ondulatoria de las capas superiores de nubes de unos 100 a 200 kilómetros de espesor. Cuando dos modones colisionan, se fusionan y expulsan materia por uno de sus bordes; este fenómeno se ha observado en la interacción de algunos vórtices de Júpiter. Surge, entonces, un nuevo vórtice más robusto y estable; así, absorbiendo otros menores, quizás formados por convección, podría mantenerse durante largo tiempo frente a los efectos destructivos de la fricción.

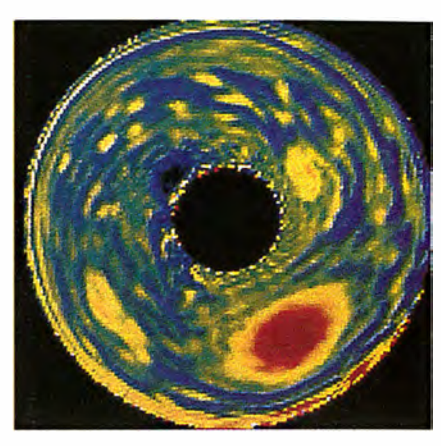
Queda mucho por averiguar sobre la posibilidad de una extensión en profundidad de las corrientes zonales; si lo hacen, habría que saber hasta dónde. T. E. Dowling, alumno de doctorado de Ingersoll, ha indagado en este problema, analizando la estructura de la región subyacente bajo las nubes a partir de los mapas de velocidades medidos en el entorno de la GRS. Sus investigaciones indican, por de pronto, que la región inferior puede desempeñar un papel decisivo en la configuración de los vórtices. Además, con un nuevo modelo matemático que incluye los efectos de esa capa, ha logrado reproducir vórtices aislados y estables, que se nutren de otros formados por inestabilidades en la cizalla meridional del viento zonal.

Recientemente, y en la línea de los modelos anteriores, Philip S. Marcus, de la Universidad de Berkeley, ha presentado nuevas simulaciones numéricas de las ecuaciones dinámicas bajo distintas condiciones reinantes en la capa inferior. En la zona caótica de fuerte cizalla del viento, entre las corrientes alternantes, se forman espontáneamente remolinos pequeños, de modo que, si su rotación intrínseca tiene el mismo sentido que la debida a la cizalladura, tienden a fusionarse hasta crear un vórtice elíptico y estable. Por contra, si la rotación es en sentido opuesto, tienden a disgregarse. A pesar de que la vorticidad decae a causa de la fricción por los extremos Norte y Sur, por donde discurren los vientos en direcciones opuestas, el sistema se mantiene por coalescencia de vórtices pequeños. En palabras de Marcus, la Mancha Roja resulta ser el producto final de un fluido que se autoorganiza al trasferir energía e información desde las pequeñas escalas hacia las grandes. Se adquiere la impresión de que el orden en gran escala emerge espontáneamente del movimiento caótico y turbulento de pequeña escala.

Con el fin de confirmar estos resultados teóricos, J. Sommeria y sus colaboradores, de la Universidad de Texas, han llevado a cabo un experimento de laboratorio estudiando el movimiento de un fluido en un recipiente cilíndrico bajo fuerte rotación. Para evocar la cizalla del flujo zonal, bombearon fluido al tanque a lo largo de un radio, extrayéndolo por el otro. Ajustando la razón de bombeo, la corriente se vuelve inestable y se crean remolinos turbillonarios, que se fusionan si la corriente es intensa y acaban por establecer el gran vórtice de la predicción de Marcus.

¿Cuáles son las perspectivas futuras para estas investigaciones? Por una parte, gracias a las imágenes de alta resolución obtenidas con los Voyager sabemos que la formación de vórtices semejantes a la GRS es fenómeno que comparten las atmósferas de Saturno y Neptuno. Destaca, en particular, la Gran Mancha Oscura de Neptuno, formación nubosa desconcertante, porque, a grandes similitudes con la Mancha Roja (tamaño relativo al del planeta, circulación anticiclónica, localización en latitud y otros), se unen importantes diferencias (alta velocidad zonal hacia el Oeste de unos 1100 kilómetros por hora, cambios rápidos de forma, vorticidad menor que la ambiente). Los estudios comparativos de estos vórtices, en su relación con las propiedades fundamentales de los planetas en los que se forman (por ejemplo, estructura térmica, fuentes de energía, velocidad angular de rotación y otros), junto con experimentos de dinámica de fluidos en el laboratorio que simulen las condiciones planetarias, pueden ser de gran ayuda.

Por otra parte, el refinamiento de los modelos postulados o la proposición de cualquier nueva teoría requerirán mayor información sobre la estructura de la atmósfera por debajo del nivel de las nubes, así como dilucidar cuál es la fuente de energía que mantiene a los vórtices durante largo tiempo en un ambiente turbulento. El avance en esta dirección podrá lograrse con las investigaciones que se efectúen en el marco del proyecto Galileo: un vehículo orbital que estudiará Júpiter durante al menos dos años, y una sonda que descenderá en su atmósfera, explorando *in situ* su estructura vertical. Lanzada con éxito el 18 de octubre de 1989, se prevé su llegada al planeta hacia 1995. Los estudios complementarios con los telescopios permitirán conocer mejor la actividad atmosférica.







9. MODELO DE MANCHA ROJA según los cálculos de ordenador efectuados por P. S. Marcus (*arriba*), comprobados en el laboratorio por J. Sommeria y sus colaboradores (*centro*). Se postula que el vórtice es la consecuencia de la coalescencia de pequeños remolinos que se forman en un ambiente caótico originado por la fuerte cizalla del viento. Si su sentido de rotación coincide con el de la cizalla, su fusión da lugar a un enorme vórtice elíptico y coherente, reminiscente de la GRS. Compárese con la imagen del Voyager que muestra el hemisferio Sur de Júpiter a partir de una proyección polar (*abajo*).

BIBLIOGRAFIA COMPLEMENTARIA

VOYAGE TO JUPITER. D. Morrison y J. Samz. NASA SP-439. 1980.

JÚPITER. G. Hunt y P. Moore. Mitchell Beazley Publishers; Londres, 1981.

EL SISTEMA SOLAR. R. Canal (comp.). Prensa Científica; Barcelona, 1987.

PLANETAS EXTERIORES. A. Sánchez Lavega y M. Cruz. Madrid, 1987.

Pirotecnia

Los secretos que se esconden tras los deslumbrantes fuegos artificiales no escapan a la curiosidad científica. Principios similares se utilizan en dispositivos tan diferentes como las cerillas y la lanzadera espacial

John A. Conkling

Tras un estallido distante y sordo, se forma en el cielo nocturno una estela de chispas de color amarillo anaranjado, que culminan en una explosión de brillantes luces azules y verdes. Otra explosión crea un arco rasgado de guirnaldas rojas, seguidas por una cascada de chispas blancas y doradas. Una tercera carcasa produce una cortina de destellos muy brillantes de luz blanca y un ruido ensordecedor.

Todos estos efectos son habituales en las grandes celebraciones desde hace siglos. Durante la mayor parte de este tiempo, el diseño y composición de los fuegos artificiales ha sido una técnica artesanal, no una ciencia. Sólo durante los últimos decenios han comenzado los investigadores a descubrir los procesos físicos que gobiernan la producción de los brillantes colores y los efectos especiales que los caracterizan. Como resultado de estas investigaciones, ha surgido la verdadera disciplina de la pirotecnia, "la ciencia del fuego". La moderna pirotecnia se ocupa no sólo de los fuegos artificiales, sino también de una batería de dispositivos que utilizan materiales semejantes, entre los que se incluyen los cohetes de señales, las cerillas corrientes e incluso los productos empleados como combustible sólido en la lanzadera espacial.

Uno de los compuestos pirotécnicos más antiguos, la pólvora negra, sirve de propulsor y explosivo en las carcasas de los fuegos artificiales modernos. La pólvora negra (o antigua pólvora de artillería) fue inventada por los chinos hace más de mil años y usada en la fabricación de toscos cohetes y petardos. Este invento llegó al mundo occidental durante la Edad Media. En 1242 el monje inglés Roger Bacon reveló la fórmula de una mezcla explosiva en el curso de su defensa contra las acusaciones de brujería. Bacon consideró que este producto era tan peligroso que escribió dicha fórmula en clave. Cuando se difundió el conocimiento de la fórmula, la pólvora negra revolucionó la industria de las canteras y la construcción. Los mosquetes y cañones que se desarrollaron a lo largo del siglo XIV utilizaban la pólvora negra como propulsor.

JOHN A. CONKLING es director ejecutivo de la Asociación Americana de Pirotecnia y profesor de química en el Colegio Universitario Washington de Chesterton, Maryland, donde imparte todos los veranos varios seminarios sobre pirotecnia; éstos son los únicos cursos sobre el tema que existen actualmente en los Estados Unidos. En 1969 se doctoró en química por la Universidad Johns Hopkins. Es autor de muchas publicaciones sobre la química de la pirotecnia y la seguridad de los fuegos artificiales.

La fórmula básica de la pólvora negra se ha mantenido prácticamente inalterada a lo largo de los siglos: una mezcla íntima de nitrato potásico (conocido popularmente con el nombre de salitre), carbón vegetal y azufre, en una proporción en peso del 75:15:10. En realidad, tal vez sea éste el único producto químico que se prepara hoy utilizando los mismos ingredientes, las mismas proporciones y el mismo proceso de fabricación que se usaban en tiempos de Colón. Esta constancia refleja el hecho de que la pólvora negra es un material pirotécnico casi ideal. Está compuesto por productos químicos abundantes y baratos, de baja toxicidad y ambientalmente seguros. La mezcla es tan estable que puede almacenarse durante decenas de años sin sufrir ningún deterioro, siempre que se mantenga seca. La pólvora negra arde con facilidad mediante una aportación moderada de energía, como la proporcionada por una chispa o una mecha encendida.

A lo largo de la historia, sólo un puñado de familias han dominado la industria de los fuegos artificiales en el mundo occidental. Detalles tales como la composición química o los procedimientos de mezcla se guardaron en secreto durante mucho tiempo, pasando de una generación a otra. Las familias han mantenido su predominio en las industrias relacionadas con todos estos artículos. Por ejemplo, en España tenemos los Caballer y, en los Estados Unidos, los Grucci o los Zambelli. Una conse-



cuencia del secreto familiar es que, hasta hace pocos decenios, la investigación pirotécnica básica apenas ha existido, y, cuando la había, sus resultados no solían aparecer en las revistas científicas.

En principio, un proceso pirotécnico no se diferencia de una combustión ordinaria. Los compuestos pirotécnicos están formados básicamente por una fuente de oxígeno (oxidante) y un agente reductor (combustible). Estos dos componentes suelen presentarse en forma de productos químicos sólidos y separados, que deben mezclarse físicamente. Cuando se aplica calor, se produce una reacción de transferencia de electrones o reacción de oxidación-reducción (redox).

Los átomos del combustible ceden electrones a los átomos del oxidante. En virtud de ese proceso, los átomos del combustible se enlazan con los átomos de oxígeno que se liberan del oxidante y forman un producto de reacción estable. Los nuevos enlaces químicos son más estables que los originales y, por tanto, se libera energía en forma de calor; en la combustión ordinaria tiene lugar el mismo proceso. Sin embargo, en el caso de la combustión el oxígeno procede del aire. En las mezclas pirotécnicas, el oxígeno forma parte del producto y el calor liberado queda mucho más confinado.

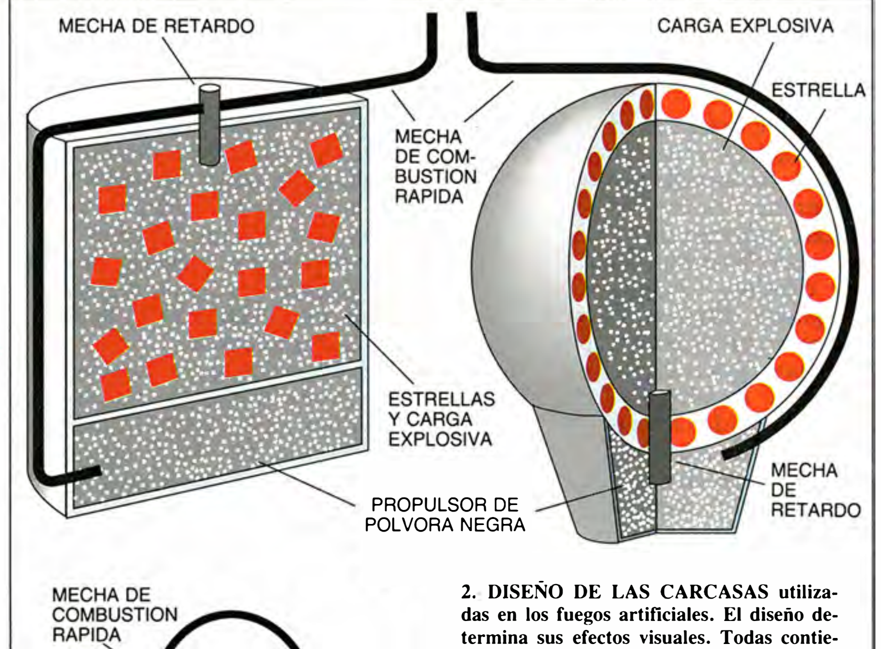
Siempre que se mantengan a bajas temperaturas y en ausencia de humedad, las mezclas pirotécnicas suelen ser bastante estables. Las mezclas sólidas experimentan sólo una reacción superficial muy lenta, controlada por difusión. Cuando se quema el compuesto, empieza a fundirse y vaporizarse, originando la llamarada pirotécnica, y los combustibles y oxidantes se mezclan de forma íntima. Esta proximidad hace que las reacciones químicas subsiguientes sean muy rápidas y, con ello, se produce una liberación de energía todavía más rápida.

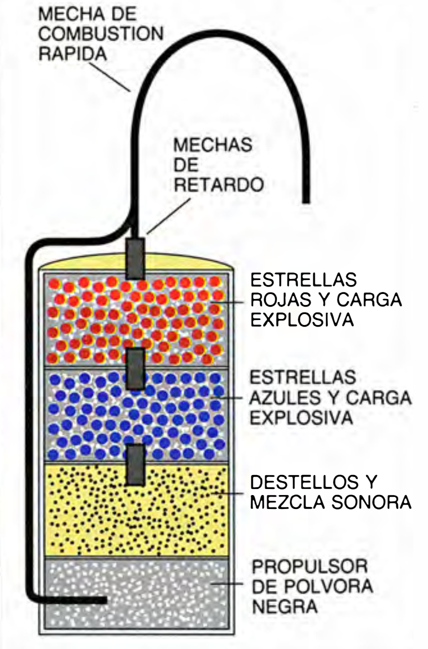
Los pirotécnicos utilizan una amplia gama de materiales combustibles. Muchas de las mezclas empleadas están constituidas por materiales orgánicos (que contienen carbono), como el carbón vegetal (en los fuegos artificiales y la pólvora de artillería) o el azúcar (en las granadas de humo). Otros combustibles corrientes contienen elementos no metálicos, como el azufre, el silicio y el boro. El silicio y el boro liberan una gran cantidad de energía cuando se oxidan y, al hacerlo, no producen gases. Se utilizan en la fabricación de mechas lentas para quemar otras composiciones con un cierto retraso. Los combustibles de metales químicamente activos —por lo común, aluminio, magnesio y titanio— arden a temperaturas elevadas y emiten una luz muy brillante. Estos materiales, que se comenzaron a usar en el siglo XIX, mejoraron extraordinariamente la luminosidad de las explosiones pirotécnicas.

Los espectaculares destellos luminosos producidos en los fuegos artificiales son el fenómeno pirotécnico más conocido. El color de esta luz depende de su longitud de onda. La luz visible consiste en radiaciones elec-



1. ESPLENDOR de los fuegos artificiales, puesto de manifiesto en esta fotografía tomada en la bahía de Nueva York durante las celebraciones del centenario de la Estatua de la Libertad. Los fuegos artificiales han existido durante más de mil años, pero los productos capaces de emitir colores intensos no se desarrollaron hasta el siglo pasado. La composición y preparación de los artículos pirotécnicos ha solido mantenerse como estricto secreto familiar; sólo en épocas recientes se han empezado a descifrar los procesos químicos que se hallan tras las luces y las explosiones.





2. DISEÑO DE LAS CARCASAS utilizadas en los fuegos artificiales. El diseño determina sus efectos visuales. Todas contienen pólvora negra como propulsor y una mecha rápida para quemarla, lo que proyecta la carcasa fuera de un tubo de lanzamiento. Una mecha de retardo detona la carga cuando la carcasa ha ascendido ya hasta cierta altura sobre el suelo. La carga explosiva produce la combustión de una mezcla de pequeñas partículas (*estrellas*) de compuestos emisores de luz. Las carcasas de tipo euro-americano contienen una mezcla homogénea de estrellas y carga explosiva (*parte superior izquierda*). Producen destellos de color de distribución irregular. En las carcasas de crisantemo o de tipo japonés, las estrellas rodean a la carga (*parte superior derecha*), y las explosiones que producen son simétricas. Las carcasas múltiples constan de varios compartimentos separados, cada uno de los cuales presenta una composición diferente; los sucesivos compartimentos están unidos entre sí por mechas de retardo (*izquierda*). La fuerte explosión con que suelen terminar los cohetes está producida por un compartimento en el que estrellas y pólvora se han sustituido por polvo de aluminio.

tromagnéticas cuyas longitudes de onda están comprendidas entre 380 y 780 nanómetros (un nanómetro equivale a la mil millonésima parte de un metro). Los rayos visibles de mayor longitud de onda son de color rojo; los de menor longitud de onda, de color violeta. Un objeto incandescente parece de color blanco cuando las longitudes de onda que radia cubren todo el espectro. Si la mayoría de la luz se emite en una parte limitada del espectro, su color corresponde a dicha parte.

Los compuestos pirotécnicos emiten luz gracias a tres procesos fundamentales: incandescencia (radiación de cuerpo negro), emisión atómica y emisión molecular. La incandescencia se produce cuando las partículas sólidas o líquidas existentes en la llama pirotécnica se calientan a temperaturas muy elevadas. Las partículas calientes emiten un espectro amplio de radiación al liberarse de su energía sobrante. La longitud de onda en la que se emite la mayor parte de la radiación disminuye a medida que aumenta la temperatura de dichas partículas. La intensidad de la emisión es proporcional a la cuarta potencia de la temperatura de la llama, de forma que un pequeño aumento de la temperatura se traduce en un aumento considerable de su brillo.

El combustible de los destellos de luz blanca es un metal reactivo (como el magnesio). Las partículas de óxido metálico sólido, formadas en la oxidación del combustible, se calientan hasta temperaturas por encima de los 3000 grados Celsius; a estas temperaturas, su color incandescente corresponde al rojo vivo. Una mezcla de perclorato potásico y polvo fino de aluminio o de magnesio produce una potente explosión y una llamarada de luz blanca. Estas mezclas de "flash fotográfico" o de "luz y sonido" tienen muchas aplicaciones, desde la fabricación de artículos pirotécnicos hasta la producción de efectos especiales en conciertos de rock o de fogonazos luminosos en fotografía nocturna. Son ellas las que originan el brillante destello en que suele terminar una explosión pirotécnica.

Las partículas metálicas de mayor tamaño mantienen su calor durante más tiempo que los polvos finos y pueden proseguir su combustión tomando el oxígeno del aire. Estas partículas producen chispas blancas, en vez de un fogonazo instantáneo. Cuanto mayores son las partículas, más duran las chispas. Las partículas de carbón vegetal y de hierro no se calientan hasta temperaturas tan altas como las partículas de metales activos —sólo se alcanzan unos 1500 grados Celsius— y, por tanto, las chispas que producen, de color dorado, son menos brillantes.

Los vivos colores que suelen contemplarse en los modernos fuegos artificiales tienen su origen en los átomos o las moléculas existentes en forma de vapor en las llamas pirotécnicas. En el primer caso, el calor generado por la llama excita uno de los electrones del átomo en cuestión y lo lanza desde el estado fundamental de su orbital normal hacia un estado de mayor energía. El electrón regresa rápidamente a su estado fundamental y emite la energía excedente en forma de un fotón (una partícula individual o unidad de radiación) con una determinada longitud de onda.

El sodio constituye uno de los más potentes emisores de luz atómica. Los átomos de sodio calentados por encima de los 1800 grados Celsius producen una luz de color amarillo-naranja cuya longitud de onda es de 589 nanómetros. Por su eficacia, este proceso sobresale frente a cualquier otra fuente de luz atómica o molecular presente en una llama pirotécnica. Bastan pequeñas cantidades de impurezas que contengan sodio para arruinar los esfuerzos realizados por los pirotécnicos en producir llamas de cualquier otro color.

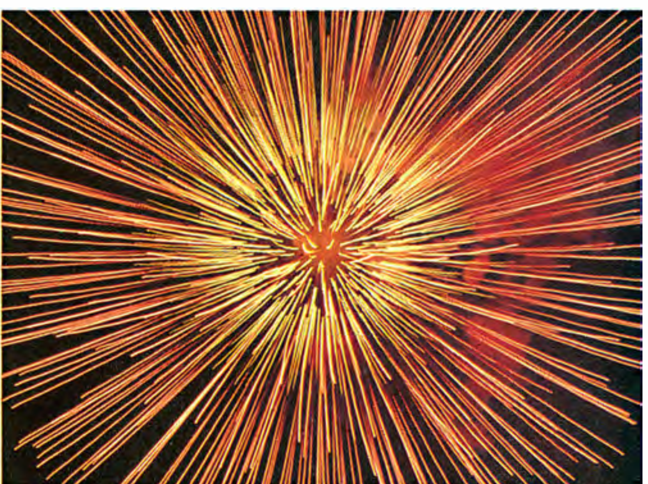
En otras aplicaciones, la prodigiosa emisión de luz por parte del sodio puede resultar de gran ayuda. La combinación de nitrato sódico, un

compuesto oxidante, con magnesio metálico produce un combustible que se utiliza mucho por los ejércitos para iluminar las operaciones nocturnas. El nitrato sódico oxida el magnesio cuando la mezcla en cuestión se quema; las partículas de óxido de magnesio resultantes emiten una luz incandescente de color blanco brillante. Las altas temperaturas (3600 grados Celsius) que se producen en las llamas de magnesio ensanchan el intervalo de longitudes de onda emitidas por los átomos de sodio. El resultado es una luz blanca intensa.

Al igual que sucede con la emisión atómica, en la emisión molecular se produce una transición desde un estado fundamental hasta un estado excitado. En las llamas pirotécnicas, las moléculas implicadas deben estar en forma gaseosa, y deben calentarse hasta una temperatura lo suficientemente elevada como para alcanzar el estado excitado que produce la radiación. Ahora bien, si la llama está demasiado caliente, la molécula se desintegra en sus átomos componentes y no se emite luz. Por otra parte, para producir luz de colores intensos, las moléculas tienen que hallarse suficientemente concentradas en la llama y la producción de partículas sólidas o líquidas debe ser mínima, dado que estas partículas emitirían una radiación incandescente que diluiría un tanto los colores.

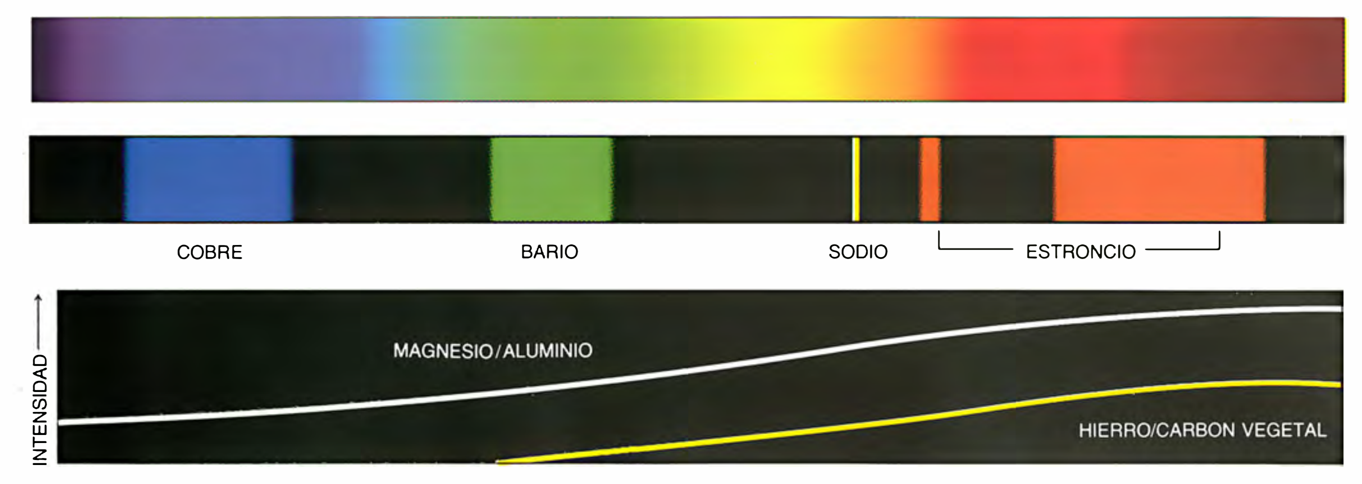
En ausencia de cualquier base teórica, la producción de colores en los fuegos artificiales fue el resultado de muchos ensayos y pruebas. Durante los últimos decenios, Bernard E. Douda y Henry A. Webster III, del Centro Logístico de la Armada en







3. DIVERSIDAD DE FORMAS Y COLORES creadas en los fuegos artificiales. Revelan los tipos de carcasas utilizados y la composición de las estrellas que contienen. Un gran número de rayos luminosos desiguales indica la utilización de una carcasa de tipo euro-americano (*parte superior izquierda*). Las explosiones de forma regular y redonda, en configuración floral, están producidas por las bien denominadas carcasas de crisantemo (*parte superior derecha*). Las carcasas de explosión múltiple o las andanadas de carcasas de explosión simple producen efectos espectaculares cuando emplean diversos compuestos pirotécnicos, capaces de emitir una amplia gama de colores y sonidos diferentes al quemarse (*fotografía inferior*).



4. LOS PRODUCTOS PIROTECNICOS producen luz por emisión incandescente, atómica o molecular. El color depende de la región del espectro visible en que se sitúa dicha emisión (*parte superior*). Los átomos y las moléculas se excitan a las altas temperaturas que se generan en una llama pirotécnica y liberan su exceso de energía en forma de luz. El sodio atómico es un emisor de color amarillo-naranja intenso. Los compuestos inestables que contienen cobre, bario y estroncio producen colores azul, verde y rojo, respectivamente (*centro*). Estas moléculas inestables se forman durante un tiempo muy corto en la llama caliente de una explosión. Las partículas de aluminio o magnesio producen chispas incandescentes de color blanco cuando se queman. Las partículas de hierro o carbón vegetal no alcanzan temperaturas tan elevadas y la luz que irradian es de color dorado.

Crane, Indiana, y David R. Dillehay, de la Longhorn Division de la Thiokol Corporation en Marshall, Texas, han llevado a cabo una serie de investigaciones para identificar los principales emisores de colores en pirotecnia. Takeo Shimizu, de la empresa pirotécnica japonesa Koa, también ha hecho contribuciones a este campo.

Unos pocos grupos de moléculas son responsables de casi todos los colores que se producen en los fuegos artificiales. Los compuestos del elemento estroncio producen los rojos: el hidróxido de estroncio (SrOH) y el cloruro de estroncio (SrCl) emiten luz roja con longitudes de onda comprendidas entre 605 y 682 nanómetros. Las moléculas que contienen bario producen los verdes. Por ejemplo, el cloruro de bario (BaCl) emite luz verde con longitudes de onda entre 507 y 532 nanómetros.

Estas moléculas, muy frágiles, son inestables a temperatura ambiente; no pueden, pues, utilizarse directamente en la fabricación de artículos pirotécnicos. Por ello se las sintetiza en el curso de reacciones rápidas que tienen lugar en la llama. Los fabricantes emplean caucho clorado, cloruro de polivinilo (un plástico que contiene cloro) y oxidantes de clorato o perclorato (que tienen un átomo de cloro con tres o cuatro átomos de oxígeno, respectivamente). Todos estos compuestos se descomponen a temperaturas elevadas y producen cloro libre. Los átomos de cloro se combinan entonces con el bario o el estroncio y crean las fugaces y deseadas moléculas productoras de luz.

La producción de una llama de color azul intenso es quizás el máximo de los retos para los pirotécnicos. El mejor emisor de azul que se ha encontrado hasta ahora, el cloruro de cobre (CuCl), es inestable a las altas temperaturas que se necesita alcanzar para producir la luz intensa de los fuegos artificiales. Cuando la temperatura de la llama supera la necesaria para que se produzca una emisión molecular óptima, las moléculas se desintegran rápidamente. En consecuencia, la obtención de fuegos artificiales de color azul exige un control muy preciso de las proporciones relativas y del tamaño de las partículas de los productos químicos empleados. Esto es válido también para los colores púrpura y violeta, creados por la emisión combinada del cloruro de estroncio y del cloruro de cobre formados en la llama. Yo pongo mucha atención a los colores de las llamas cuando contemplo unos fuegos artificiales; si aparece un color azul aceptable, me impresiona y me entra la curiosidad por saber la mezcla química utilizada.

La combinación de los compuestos que producen los colores con los oxidantes y combustibles adecuados puede dar lugar a ciertos efectos especiales. El color rojo de determinadas chispas se debe a una combinación de carbonato de estroncio (que emite luz roja) y gránulos de aluminio (que producen las chispas). Estos dos ingredientes se mezclan con combustibles, aglomerantes y un oxidante para crear una pasta espesa; si se preparan bengalas, en esta mezcla se introducen alambres y, finalmente, el producto se deja secar. Otro compuesto de estroncio, el nitrato de estroncio, se mezcla con perclorato potásico (un oxidante y fuente de cloro) y varios combustibles para producir las bengalas rojas que se utilizan como señales de peligro en las carreteras.

La estructura de un cohete de fuegos artificiales es también el resultado de una complicada mezcla de artesanía e ingeniería. Hay dos tipos de "carcasas" pirotécnicas. Las carcasas cilíndricas utilizadas en América y Europa, con un diámetro comprendido entre los 7 y 30 centímetros, se disparan mediante tubos de mortero metálicos, de cartón o de plástico. La combustión de cierta cantidad de pólvora negra situada en su parte inferior impulsa la carcasa por el aire hasta una altura de algunos centenares de metros. Cuando la carcasa inicia su ascenso, empieza a quemarse una mecha de retardo; unos segundos más tarde, cuando ya se encuentra a determinada altura sobre el suelo, una carga de pólvora negra rompe su cubierta y quema los gránulos del compuesto que produce los colores (las llamadas estrellas), que se encuentran distribuidos irregularmente por el interior. Las estrellas se difunden formando una pauta caprichosa de luz y color. Existe otro tipo de carcasas, conocido con el nombre de 'saludo', que no contiene estrellas, sino otros compuestos que producen simplemente un fogonazo luminoso y una potente explosión.

Las carcasas japonesas de crisantemo poseen un diámetro similar a las europeas y americanas; se lanzan también con ayuda de morteros. Sin embargo, en las carcasas de crisantemo, las estrellas antes citadas están

dispuestas esféricamente alrededor de una carga explosiva central de pólvora negra. Cuando la carga explota, las numerosas estrellas contenidas en la carcasa arden y se distribuyen formando una figura redonda y simétrica. Dependiendo del tamaño y de la composición química de las estrellas, el resultado puede variar desde un destello rápido hasta una cola alargada. Esta cola puede incluso ir cambiando de color, si las estrellas constan de más de una capa de sustancias productoras de colores.

Algunas carcasas del tipo euroamericano contienen varios compartimentos, cada uno de los cuales lleva su propia carga explosiva y de estrellas (o de pólvora de fogonazo y ruido). Cuando uno de estos compartimentos explota, prende una mecha de retardo que conduce al siguiente. De esta forma, una simple carcasa produce una serie de explosiones múltiples. Resulta increíble que las barreras protectoras que separan los sucesivos compartimentos explosivos estén fabricadas con un material tan poco exótico como el cartón.

Además de con la luz, la pirotecnia se relaciona frecuentemente con la producción de calor. Los artículos pirotécnicos más conocidos para la producción de calor —las cerillas corrientes— están constituidos por una mezcla de clorato potásico oxidante y de azufre con un material combustible y aglomerante.

La combustión de una mezcla de siliciuro de calcio y óxido de hierro desprende una cantidad moderada de calor y no produce ningún gas. Durante la segunda guerra mundial, se fabricaron unos pequeños dispositivos pirotécnicos que contenían esta mezcla y una mecha, incorporándolos a unos botes especiales de comida de campaña para calentar los alimentos envasados sin necesidad de encender fuego. La combustión retardada de ciertas mezclas, que se presentan en forma de barras con un combustible de boro, tungsteno o silicio, produce una emisión controlada de calor durante un período de tiempo determinado, lo que provoca luego otra reacción mayor. Estas mezclas se utilizan para controlar la secuencia temporal de algunos dispositivos especiales de uso en técnica aeroespacial, entre los que se pueden citar las tuercas explosivas que hacen funcionar con rapidez las salidas de emergencia o desprenden los cuerpos ya utilizados de los cohetes espaciales; dispositivos de retardo semejantes impiden que las granadas de mano exploten inmediatamente después de extraer el pasador de seguridad y liberar la correspondiente anilla. Se han desarrollado también algunos productos pirotécnicos que sirven de señuelos para proteger a los aviones de los cohetes enemigos guiados por calor. Estos compuestos emiten radiación infrarroja que simula la emisión térmica del motor de un avión a reacción.

La producción de calor suele ir asociada con la emisión de humo y gas. Las granadas de humo de colores, utilizadas en dispositivos de señalización y en los fuegos artificiales diurnos, están constituidas por una mezcla de clorato potásico, como oxidante, y de azúcar, como combustible, que, cuando se activa, produce la vaporización



5. UNA CERILLA O FOSFORO CORRIENTE es un dispositivo pirotécnico singular. La reacción entre el clorato potásico oxidante que hay en la cabeza y el fósforo rojo que hay en el rascador origina una llama, que quema la mezcla de clorato potásico y aglutinante combustible en que consiste la cabeza de la cerilla. Todos los efectos pirotécnicos se encuentran presentes: calor, humo, luz, gas y sonido. Algunas cerillas antiguas, hechas de clorato potásico, se encendían empapándolas en ácido sulfúrico. La combinación del azufre y del clorato potásico resulta tan sensible a la ignición que las mezclas de estos dos materiales fueron prohibidas por la ley en Inglaterra en el año 1875. El uso de clorato potásico en los fuegos artificiales está restringida en Estados Unidos y otros países.





6. LOS FUEGOS ARTIFICIALES han alegrado las fiestas durante siglos. Una cromolitografía de la inauguración en 1883 del Puente de Brooklyn sobre el East River recoge una imagen particularmente sugerente (*arriba*). El centenario de este mismo puente fue conmemorado con otra extraordinaria demostración del arte pirotécnico (*abajo*). El disparo de fuegos artificiales de 1983 fue obra de los Grucci, una de las familias más importantes dedicadas a la pirotecnia en los Estados Unidos.

de ciertos pigmentos orgánicos que crean nubes de humo fuertemente coloreadas. Se recurre al azúcar porque arde a bajas temperaturas; la creación de una llama más caliente desintegraría los pigmentos.

Los cohetes de combustible sólido son, en esencia, unos dispositivos pirotécnicos de enorme tamaño diseñados para optimizar la producción de gases. Cada cohete de propulsión de la lanzadera contiene medio millón de kilos de un material propulsor consistente en un combustible de aluminio pulverizado y un oxidante de perclorato amónico; la mezcla incluye también un combustible y aglomerante especial conocido como PBAN, un polímero de polibutadieno, ácido acrílico y acrilonitrilo. Cuando se oxida, el PBAN libera enormes cantidades de monóxido y dióxido de carbono gaseosos y de vapor, que ayudan a impulsar la lanzadera hacia el espacio. El perclorato amónico es el material adecuado para esta aplicación, ya que todos sus productos de descomposición son gaseosos, circunstancia que favorece la impulsión de los cohetes.

La producción de gases en menor escala origina los silbidos que se oyen en algunos castillos de fuegos artificiales. Ciertas mezclas de perclorato potásico como oxidante y una sal de ácido orgánico (como el salicilato de sodio) son capaces de arder capa a capa y producen emisiones intermitentes de gas. Comprimidas en unos tubos de pequeño diámetro, los rápidos chorros sucesivos del gas que emiten producen silbidos.

La aplicación más adecuada para una mezcla pirotécnica concreta viene determinada, en gran medida, por la reactividad de su oxidante y su combustible. La reactividad de los combustibles guarda una estrecha relación con la cantidad de energía (calor de combustión) que se libera cuando se combinan con oxígeno. Los metales liberan cantidades muy grandes de energía al oxidarse; el azúcar, relativamente poca. El carbón vegetal y otros productos naturales como la goma roja, que es una resina, generan la cantidad de calor moderada que se necesita para activar los compuestos que originarán el color de los fuegos artificiales.

La reactividad de los oxidantes depende de dos factores principales: la temperatura de descomposición y el calor de descomposición. Cuando se alcanza la temperatura de descomposición, los oxidantes empiezan a liberar oxígeno en cantidades significativas. El calor de descomposición, como indica el término, expresa la cantidad de calor que se necesita para descomponer un oxidante determinado con el fin de liberar oxígeno. Esta cantidad puede ser positiva (reacción endotérmica), en cuyo caso la descomposición absorbe calor, o negativa (reacción exotérmica), en cuyo caso lo genera.

El clorato potásico se descompone ya a una temperatura de 360 grados Celsius y la reacción es exotérmica; se utiliza en la fabricación de granadas de humo y de cerillas corrientes debido precisamente a que dicha reacción es exotérmica y se pone en marcha con facilidad. En el extremo opuesto, el óxido de hierro se descompone a 1500 grados Celsius y la reacción es fuertemente endotérmica; la descomposición de este oxidante sólo puede producirse mediante un combustible metálico muy activo, como es el aluminio.

La forma de empaquetarla y la homogeneidad de una mezcla pirotécnica pueden afectar también a su velocidad de reacción. Como bien saben todos los fabricantes de estos dispositivos, el confinamiento compacto de un material pirotécnico se traduce en una concentración de calor y gases de combustión en las proximidades de la zona de reacción. Una mezcla que arde de modo progresivo en un espacio abierto puede explotar violentamente cuando se encierra en un espacio reducido. En general, cuanto mayor sea la homogeneidad del combustible y del oxidante, tanto más rápida será la velocidad de combustión.

Una vez, en un seminario que daba sobre este tema, se me preguntó por

qué no suelen utilizarse líquidos en pirotecnia, dado que se les puede mezclar mucho mejor y, por tanto, producirían compuestos más reactivos que los sólidos. La respuesta a tal pregunta es que los líquidos se mezclan demasiado bien. Los productos líquidos serían extremadamente homogéneos y, por ende, resultarían muy reactivos y muy sensibles a la ignición. Además, los líquidos podrían producir sedimentos durante su almacenamiento, con lo que se alteraría su equilibrio químico. Las primeras versiones de la dinamita, consistentes en una mezcla de materiales porosos (como el serrín de madera) embebidos de nitroglicerina líquida, eran muy inestables por esa misma razón.

La reactividad es máxima cuando el oxidante y el combustible están mezclados a nivel atómico y los átomos del oxidante, aceptantes de electrones, están situados en la proximidad inmediata de los átomos o iones del combustible, que ceden estos electrones cuando se inicia una reacción pirotécnica. Aunque estas mezclas de alta energía son, hablando en rigor, materiales explosivos y no artículos pirotécnicos, los principios que gobiernan su comportamiento son semejantes. Por ejemplo, la nitroglicerina tiene la fórmula molecular $C_3N_3H_5O_9$. Una pequeña perturbación (el calor o un impacto) produce su descomposición en dióxido de carbono (CO_2), agua (H_2O), nitrógeno (N_2) y un ligero exceso de oxígeno (O_2). En este proceso, los enlaces atómicos nitrógeno-oxígeno se substituyen por enlaces carbono-oxígeno, hidrógeno-oxígeno y nitrógeno-nitrógeno, mucho más estables que aquéllos; el resultado global es una violenta liberación de energía.

Un material de este tipo, menos conocido, pero cada vez más importante, es la azida de sodio, el componente activo utilizado en las bolsas de aire autoinflables. Está constituido por una red de iones de sodio y azidas (un grupo de tres átomos de nitrógeno enlazados químicamente). Un impacto violento rompe esta estructura reticular. A continuación, el sodio se combina con el oxígeno y los átomos de nitrógeno se reagrupan en parejas produciendo grandes cantidades de gas nitrógeno.

La historia de la industria pirotécnica de todos los países está llena de accidentes trágicos producidos durante los procesos de fabricación. La mejora en la seguridad exige un conocimiento detallado del fenómeno de la ignición.

La ignición se inicia cuando la energía procedente de alguna fuente —llama, fricción, impacto, chispa, temperatura elevada o incluso un haz de láser— rompe los enlaces químicos de una mezcla pirotécnica. Como resultado de este proceso, se forman enlaces más estables que los originales y se libera cierta cantidad de energía. Si esta energía logra activar la capa siguiente de la mezcla, la reacción prosigue; pero si la energía es absorbida por el material que rodea a la mezcla o resulta insuficiente para activar la capa siguiente, se extingue la reacción.

Fred L. McIntyre y sus colaboradores del Laboratorio de Ensayo de Riesgos del Centro Espacial John C. Stennis, en Mississippi, han analizado una serie de compuestos pirotécnicos para determinar su sensibilidad a la ignición utilizando diferentes fuentes de energía. Sus estudios han demostrado que los factores críticos que controlan dicha sensibilidad son la cantidad de calor producido por la reacción y la temperatura de ignición, es decir, la temperatura mínima necesaria para producir una reacción rápida y automantenida. La sensibilidad a la ignición se ve afectada también por el tamaño de las partículas de los productos químicos que componen la mezcla y por el tamaño del grano del producto resultante; los granos finos se queman con mayor rapidez que los gruesos. La facilidad de la ignición por fricción depende de la presencia de materiales abrasivos en la mezcla. La adición de un producto lubricante como la cera puede reducir la probabilidad de que se produzca una ignición por fricción.

El aspecto de la seguridad es también muy importante para los usuarios de productos pirotécnicos, que con frecuencia son niños que celebran efemérides en las que tirar cohetes resulta tradicional (las verbenas o las fallas de Valencia). Los países de la Comunidad Europea cuentan con una amplia serie de normas de seguridad, que ahora se está intentando refundir en un código único dentro de su actual proceso de unificación legal y económica. En los Estados Unidos, la Comisión de Seguridad de Productos de Consumo promulgó en 1976 unas normas federales muy estrictas para los artículos pirotécnicos usados por el público.

He podido comprobar, en numerosas conversaciones mantenidas con investigadores de muy diversos campos, que muchas carreras científicas se iniciaron bajo el estímulo de juveniles experimentos pirotécnicos. La necesidad de proporcionar a los padres una explicación rápida de por qué estaba lleno de humo un sótano, consecuencia del éxito de algún experimento, ha sido también probable detonante de más de una carrera jurídica.

El carácter de negocio familiar de la industria pirotécnica civil y la naturaleza secreta de una gran parte de los trabajos relacionados con la defensa militar hacen que la formación académica en este campo sea difícil de adquirir. Los únicos cursos regulares sobre pirotécnica que se imparten anualmente en los Estados Unidos son los seminarios de una semana de duración que se desarrollan en el Colegio Universitario Washington de Maryland, a cargo mío y de algunos colegas. Afortunadamente, son muchas las organizaciones que consideran que la pirotecnia es una ciencia y que publican los resultados de la investigación que se realiza. Entre estas organizaciones cabe citar a la Sociedad Pirotécnica Internacional y alguna otra con una orientación más dirigida a los aficionados (verbigracia, "Pyrotechnics Guild International"). Una revista no periódica que se publica con el nombre de *Pyrotechnica* contiene también artículos actuales sobre el tema.

El apoyo más sólido de la industria civil sigue siendo el favor del público espectador. Por ejemplo, en los Estados Unidos de América, el consumo anual de artículos pirotécnicos por individuo con motivo de fiestas familiares o vecinales se ha duplicado desde 1976. Los disparos de fuegos artificiales constituyen una tradición con siglos de antigüedad y todavía resultan fascinantes en la actualidad, a pesar de la competencia que les hacen los conciertos de rock, los vídeos musicales y otras formas de espectáculos visuales y auditivos. Es evidente que la tecnología moderna no ha sido capaz de superar todavía la emoción que nos produce la contemplación del trabajo de un artista pirotécnico, cuando el cielo nocturno retumba con el estallido de los truenos y se llena de brillantes lluvias de colores.

BIBLIOGRAFIA COMPLEMENTARIA

PYROTECHNICS FROM THE VIEWPOINT OF SOLID STATE CHEMISTRY. Joseph H. McLain. The Franklin Institute Press, 1980.

THE CHEMISTRY OF PYROTECHNICS. John A. Conkling. Marcel Dekker, Inc., 1985.

FIREWORKS: THE ART, SCIENCE AND TECHNIQUE. Segunda Edición. Takeo Shimizu. Pyrotechnica Publications, 1988.

FIREWORKS: A HISTORY AND CELEBRATION. George Plimpton. Doubleday & Co., Inc., 1989.

Chancro del castaño

El hongo que asoló al castaño americano sufre, a su vez, un ataque biológico. Los fitopatólogos esperan que el nuevo parásito termine por domeñar al antiguo

Joseph R. Newhouse

Hace cien años, el castaño americano (*Castanea dentata*) podría haberse considerado el árbol americano por antonomasia. Esta magnífica especie representaba una cuarta parte de los árboles de madera noble existentes en su área natural y formaba una cadena casi ininterrumpida de bosques desde Maine hasta Georgia. Crecía más deprisa que la mayoría de los otros árboles productores de madera, alcanzando a veces una altura de 30 metros y un contorno de más de dos metros. Su madera de fibra recta se troceaba fácilmente, ardía casi sin humo y resistía la podredumbre gracias a una generosa dotación de taninos, sustancia amarga que se encuentra también en las hojas de té. Tal resistencia hizo que fuese la madera ideal para vallados, postes de telégrafos y traviesas de ferrocarril. En aquellos lugares donde el árbol era abundante, los taninos sostenían una industria floreciente. La castaña, que tenía mucho más aroma que la de otras especies de castaño, era alimento básico para muchos animales e ingrediente de muchos dulces.

En la actualidad, el castaño sobrevive principalmente como un arbusto menor, que brota del sistema radicular de árboles muertos o en declive. La mayor parte de los castaños adultos, unos tres o cuatro mil millones, murieron durante la primera mitad de este siglo debido a una enfermedad fúngica, introducida, según se piensa, alrededor de 1900 con la importación de castaños procedentes de Asia. Los robles y otros árboles de madera noble han llenado el considerable hueco dejado por la muerte del castaño.

Los continuos esfuerzos realizados para producir variedades más resistentes del árbol han sido hasta ahora decepcionantes. Sin embargo, descubrimientos de biotecnología recientes traen la esperanza de que se puedan utilizar otras armas contra el hongo. Se sabe que algunas cepas del mismo son demasiado débiles para matar a su huésped, y se ha aclarado en parte el mecanismo genético responsable de esa menor virulencia. Si los fitopatólogos consiguen transmitir esta menor virulencia a todas las demás cepas, los castaños adultos podrían adornar de nuevo el paisaje norteamericano.

El chancro se observó por primera vez en 1904 por Herman W. Merkel, del Parque Zoológico de Nueva York en el Bronx. Alarmado por la aparición de chancros (depresiones necróticas) en los árboles del parque, solicitó y obtuvo un presupuesto para podar las ramas enfermas. Sin embargo, a pesar del importante esfuerzo, la mayoría de los árboles se infectaron en el plazo de un año y murieron muchos de los ejemplares más pequeños. Los árboles mayores sobrevivieron durante un decenio, pero terminaron por sucumbir.

William A. Murrill, del Jardín Botánico de Nueva York, tras determinar que el hongo entra en el castaño a través de heridas en la corteza, llamó al patógeno *Diaporthe parasitica* en 1906. En 1912, consideraciones taxonómicas llevaron a los micólogos a rebautizarlo *Endothia parasitica*, denominación por la cual fue más conocido. En 1978, de nuevo por razones taxonómicas, el hongo recibió su nombre actual, *Cryphonectria parasitica*.

Sin el obstáculo de barreras naturales y enfrentado con una población huésped completamente indefensa y relativamente homogénea, *C. parasitica* se propagó con rapidez. En 1950 ocupaba toda el área natural del castaño [*véase la figura 2*]. ¿Cómo pudo extenderse tan rápidamente? Al igual que la mayoría de los hongos, se reproduce sexual y asexualmente. Comienza por dispersar esporas sexuales, llamadas ascosporas, a través del viento, para establecer colonias alejadas. Luego, las colonias vuelven a difundirse produciendo masas pegajosas de esporas asexuales, llamadas conidios, cuyo transporte puede realizarse por las gotas de lluvia o viajar de árbol en árbol adheridas a las patas de pájaros e insectos.

Los técnicos forestales intentaron retardar el avance de la enfermedad clareando bandas de más de un kilómetro de ancho en bosques aún no afectados por el chancro y destruyendo los árboles enfermos, pero no se consiguió el objetivo. También fracasaron los intentos por repoblar una región después de retirar los castaños muertos. Esto se debe a que el hongo puede sobrevivir a la muerte de su huésped subsistiendo como parásito benigno en otras especies arbóreas o como un saprófito sobre materia orgánica muerta.

Muchos investigadores consideraron que el castaño americano estaba sentenciado a muerte, pero la capacidad del árbol para rebrotar repetidamente de sus raíces le ha salvado de la extinción. Aunque los rebrotes terminan por ser atacados por la enfer-

JOSEPH R. NEWHOUSE estudia la biología molecular y la citología de los hongos que atacan a las plantas. Es investigador asociado de la Unidad de Investigación Fitopatológica del Departamento de Agricultura en Frederick, Maryland. Se licenció en biología en el Colegio Universitario de San Vicente, Latrobe, Pennsylvania, en 1977 y se doctoró en fitopatología por la Universidad de West Virginia en Morgantown, en 1988. En su tiempo libre se dedica a la fotografía, a descubrir las posibilidades del ordenador y al golf.

1. CASTAÑO AMERICANO, de 22 metros de altura. Este ejemplar escapó al ataque del chancro porque fue plantado en Lake County, Michigan, unos 320 kilómetros al oeste del área natural de la especie. Los escasos supervivientes constituyen una valiosa reserva de reforestación.

medad, normalmente antes de la edad reproductiva, proporcionan estacas con las que puede mantenerse una reserva en ambientes protegidos. Los bosques de castaños adultos plantados al oeste de la región natural del árbol ofrecen otro rayo de esperanza.

Como primer paso hacia una cura había que reconstruir la historia de la epidemia. Si los investigadores supiesen cuál era el huésped en el que el hongo vivía inicialmente, podrían descubrir cómo se había defendido. La primera sospecha, que el hongo había entrado en EE.UU. sobre castaños asiáticos importados, se confirmó con varios hechos. En primer lugar, en 1938, el chancro atacaba al castaño europeo (*Castanea sativa*) en regiones de Italia donde se habían establecido viveros experimentales del castaño japonés (*Castanea crenata*) diez años antes. En segundo lugar, muestras de cultivo del hongo tomadas de castaños en China, Japón, Europa y EE.UU.presentaban virulencia y morfologías similares en los cultivos artificiales. En tercer lugar, *Castanea* creció en China y Japón mucho tiempo antes de que apareciese en cualquier otro lugar del mundo, como atestiguaban los datos fósiles. Por último, la especie asiática había demostrado ser resistente al hongo. Todo ello sugiere que los árboles asiáticos coevolucionaron con el hongo del chancro y que la selección natural favoreció a los resistentes.

La enfermedad progresaba con más lentitud en Europa que en Estados Unidos, en parte debido a que la especie europea es más resistente al hongo. Además, el castaño europeo es un árbol pequeño, de huerto, que suele crecer en grupos muy separados. Las distancias, junto a barreras naturales como los Alpes, hicieron más lenta la propagación de la enfermedad. A pesar de todo, en 25 años, el árbol europeo había sido diezmado en Italia y atacado en Francia.

Y entonces, hacia 1950, apareció una señal alentadora: cerca de Génova se encontraron castaños europeos cuyos chancros habían detenido



su crecimiento [*véase la figura 4*]. Antonio Biraghi, del Instituto Forestal de Florencia, descubrió que los árboles habían aislado algunas zonas enfermas con tejido de callo, una forma de corteza. Los mismos árboles tenían otros chancros que seguían propagándose. ¿A qué se debía ello? Biraghi sugirió que algunas cepas fúngicas podrían haber perdido su virulencia. Por extraño que parezca, pasó más de una docena de años antes de que nadie intentara comprobar esta hipótesis. Finalmente, en 1964, Jean Grente, del Instituto Nacional de Investigación Agrícola de Clermont-Ferrand, confirmaba que Biraghi estaba en lo cierto.

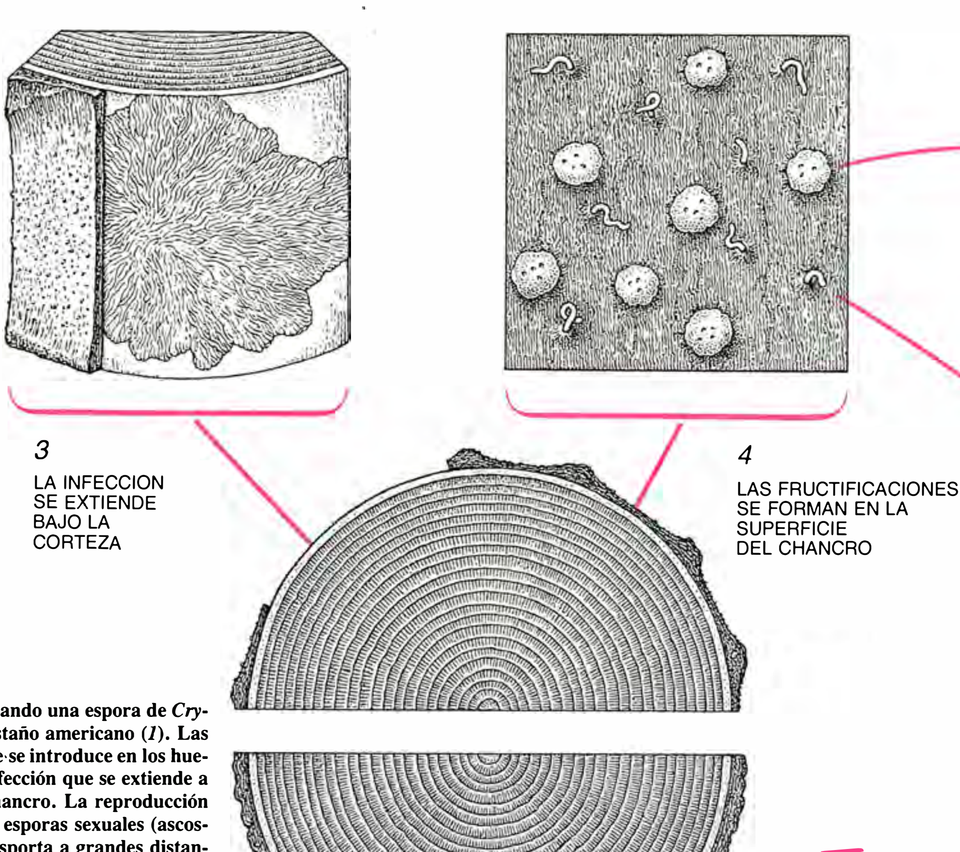
Grente extrajo el hongo de chancros confinados y observó que carecía del pigmento naranja y de la mayor parte de los picnidios, o fructificaciones asexuales, que caracterizan al hongo de los chancros en expansión. Aunque las cepas sin pigmentar podían producir chancro cuando se inoculaban en un castaño europeo, el chancro crecía tan lentamente que el árbol tenía tiempo para aislarlo. Grente llamó hipovirulentas a estas nuevas cepas, ya que no podían sostener una infección eficaz.

Lo más notable de todo fue el descubrimiento por Grente y sus colegas de que el rasgo de hipovirulencia era transmisible. Vieron que, cuando se colocaba contiguos, en placas de petri, cultivos en agar virulentos e hipovirulentos, las colonias resultantes se parecían inicialmente a aquellas de las que se habían tomado las muestras. Tras varios días, sin embargo, el resultado del crecimiento de la colonia virulenta era semejante al de la colonia hipovirulenta. Este hallazgo animó a los investigadores a tratar chancros con cepas hipovirulentas, proceso que paraba rápidamente la expansión del chancro. Estos y otros experimentos mostraron que el determinante de la hipovirulencia residía en el fluido intracelular del hongo, o citoplasma, y que se transfería de las líneas hipovirulentas a las virulentas durante la anastomosis hifal. (La anastomosis consiste en la formación de un puente entre los talos filamentosos, o hifas, de hongos próximos, lo que permite que el citoplasma fluya de una hifa a otra.)

La hipovirulencia había triunfado sobre la virulencia tanto en el laboratorio como en el campo. Grente comenzó un programa de biocontrol en Francia, en 1966, que ha tenido desde entonces un gran éxito. En Italia, por otro lado, una hipovirulencia natural ha llevado a la recuperación del castaño sin intervención humana.

Los investigadores pensaron que sería posible reproducir estos espléndidos resultados en EE.UU., y, en efecto, los primeros éxitos generaron una gran expectación. En 1972 Richard A. Jaynes, Neal K. Van Alfen y sus colegas, de la Estación experimental agrícola de Connecticut en New Haven, impidieron el crecimiento de chancros en plántulas de invernadero, inoculándolas con cepas hipovirulentas europeas suministradas por Grente. Un año más tarde, después de recibir autorización del Departamento de Agricultura de los Estados Unidos para llevar a cabo estudios de campo, inocularon cierto número de árboles con cepas europeas hipovirulentas y se detuvo igualmente el crecimiento de los chancros. Las inoculaciones habían logrado convertir la mayoría de los hongos a la forma hipovirulenta.

Sin embargo, estudios posteriores mostraron que la inoculación no siempre evitaba que un chancro siguiese creciendo. Un factor desconocido estaba impidiendo que algunas de las líneas hipovirulentas pasasen sus rasgos benignos a las cepas virulentas durante la anastomosis hifal. Tales bloqueos selectivos eran el resultado de lo que se llama incompatibilidad vegetativa. Sandra L. Anagnostakis, de la estación de Connecticut, identificó los genes que determinan la compatibilidad vegetativa en *C. parasitica*, probando luego que cuanto mayor era la variación genética entre dos cepas, menos probable era que se produjese la anastomosis hifal. Su hallazgo llevó al uso de pruebas de compatibilidad para clasificar a los hongos virulentos dentro de grupos compatibles vegetativamente, lo que permitía aplicar a los chancros la cepa hipovirulenta que se sabía experimentaba anastomosis con la raza virulenta correspondiente. Este tra-



2. El CICLO DE LA ENFERMEDAD se inicia cuando una espora de *Cryphonectria parasitica* arriba a la corteza de un castaño americano (*1*). Las esporas producen un tubo germinativo, o hifa, que se introduce en los huecos existentes en la corteza (*2*), originando una infección que se extiende a través del cambium (*3*) y que se manifiesta en chancro. La reproducción (*4-6*) se hace de dos formas, sexual y asexual: las esporas sexuales (ascosporas) son dispersadas por el viento, que las transporta a grandes distancias (*7*) y las masas de esporas asexuales (conidios) se segregan en una matriz pegajosa, que las patas de insectos y pájaros se encargan de llevar a los árboles cercanos (*8*). La lluvia puede acarrear los conidios hacia la base del árbol, produciendo múltiples infecciones. Los chancros terminan por matar al huésped, al impedir el intercambio de agua y nutrientes entre las hojas y las raíces (*9*). *Cryphonectria parasitica* tardó menos de cincuenta años en cubrir la totalidad del área natural del castaño (*mapa*).

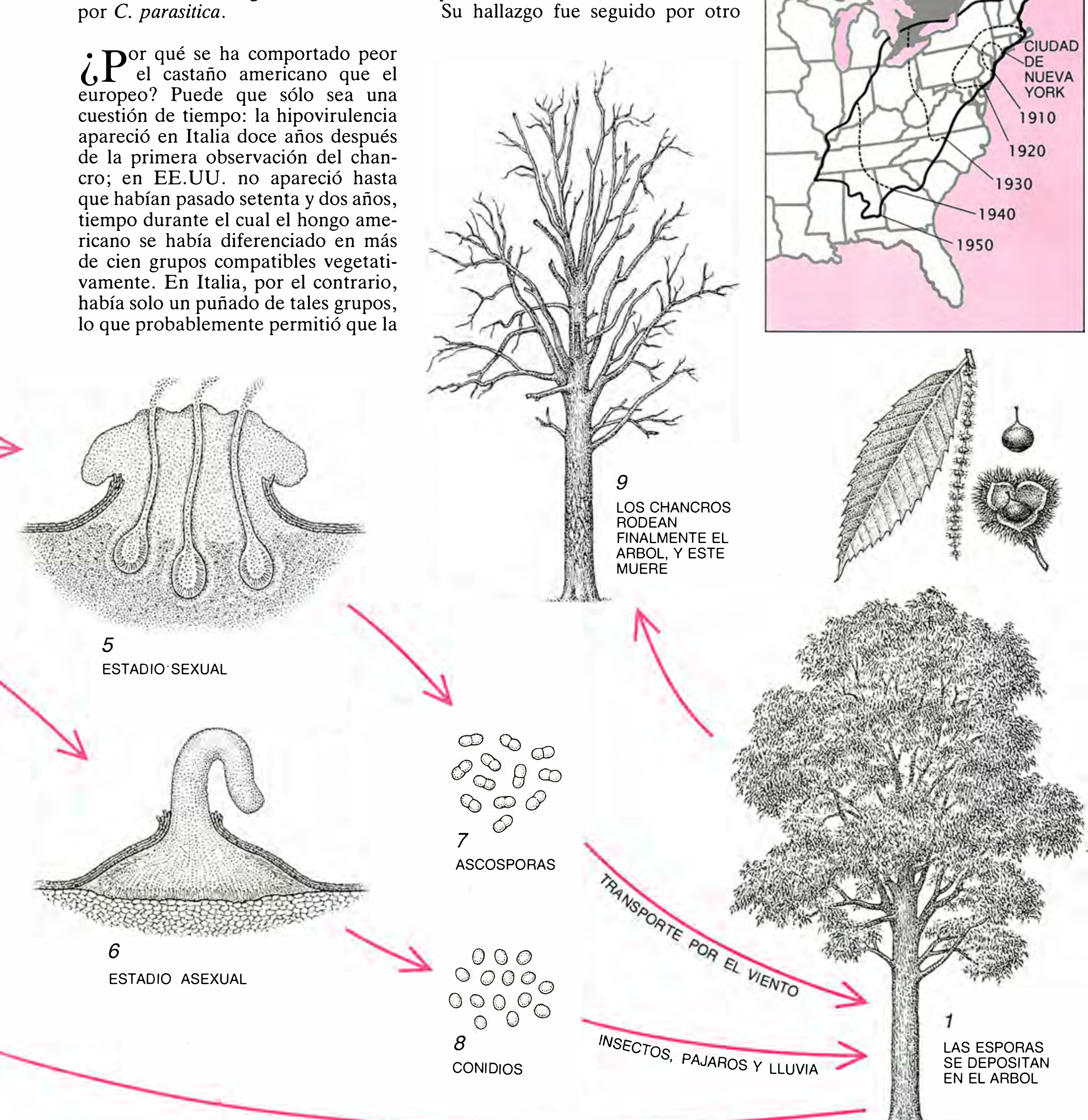
tamiento solía producir mejores resultados.

Entretanto, en 1976, se descubrieron cepas endémicas hipovirulentas en un bosque de castaños americanos al oeste de Michigan, fuera del área natural del árbol, y más tarde también en el interior de ésta. Muchos fitopatólogos se preguntaron si el castaño americano podría recuperarse por sí solo, como había hecho en Italia la especie europea. Sin embargo, ni la inoculación ni la hipovirulencia natural acabaron con la enfermedad. Ningún castaño americano ha llegado aún a adulto en regiones infestadas por *C. parasitica*.

¿Por qué se ha comportado peor el castaño americano que el europeo? Puede que sólo sea una cuestión de tiempo: la hipovirulencia apareció en Italia doce años después de la primera observación del chancro; en EE.UU. no apareció hasta que habían pasado setenta y dos años, tiempo durante el cual el hongo americano se había diferenciado en más de cien grupos compatibles vegetativamente. En Italia, por el contrario, había solo un puñado de tales grupos, lo que probablemente permitió que la hipovirulencia se extendiese más rápidamente.

A pesar de que el hongo ha eludido todos los intentos de erradicación realizados hasta ahora, descubrimientos biotecnológicos recientes sugieren que la guerra aún puede ganarse. El primer avance importante se produjo en 1975, cuando Van Alfen y sus colegas probaron que, durante la anastomosis hifal, el material del núcleo no pasa de la línea hipovirulenta a la virulenta, lo que les llevó a la conclusión de que el agente responsable de la hipovirulencia debía residir en el citoplasma.

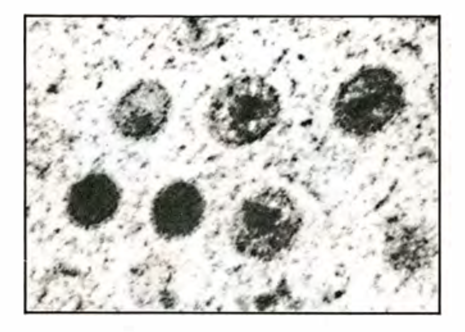
Su hallazgo fue seguido por otro descubrimiento pertinente dentro del mismo año. Richard M. Lister, de la Universidad de Purdue, y su alumna de doctorado Eileen M. Moffitt detectaron ARN bicatenario (ARNdc, ácido nucleico que codifica información genética en muchos virus fúngicos) en el citoplasma de dos cepas hipovirulentas europeas, pero no lo encontraron en las líneas virulentas ni



de Europa ni de EE.UU. Dos años más tarde, Peter R. Day y sus colaboradores, de la estación de Connecticut, confirmaron los resultados de Lister y Moffitt y encontraron ARNdc en líneas hipovirulentas norteamericanas. También mostraron que algunas cepas hipovirulentas tenían más de un segmento de ARNdc. Desde entonces, se han identificado líneas hipovirulentas que poseen hasta 10 segmentos de diferentes longitudes, pero nadie ha vinculado aún ninguno de estos segmentos con la hipovirulencia.

A pesar de ello, no cabe duda del papel del ARNdc en el origen de la hipovirulencia: la transmisión de la hipovirulencia siempre implica la transferencia de ARNdc y la eliminación del ARNdc siempre hace que las líneas hipovirulentas se conviertan en virulentas. La causa y el efecto estaban claros, pero no así el mecanismo que los relacionaba.

Como primer paso para identificarlo, los fitopatólogos buscaron los orígenes del ARNdc. Los virus eran su fuente más probable, puesto que la mayoría de los virus que se sabe que infectan hongos tienen su información genética codificada por ARNdc. En 1978, James A. Dodds, entonces en la estación de Connecticut, presentó por primera vez pruebas de la existencia de una partícula viriforme (VLP) en una cepa hipovirulenta. Las partículas que obtuvo de una línea europea hipovirulenta tenían forma de maza, de unos 100 nanómetros de longitud; poseían ARNdc.



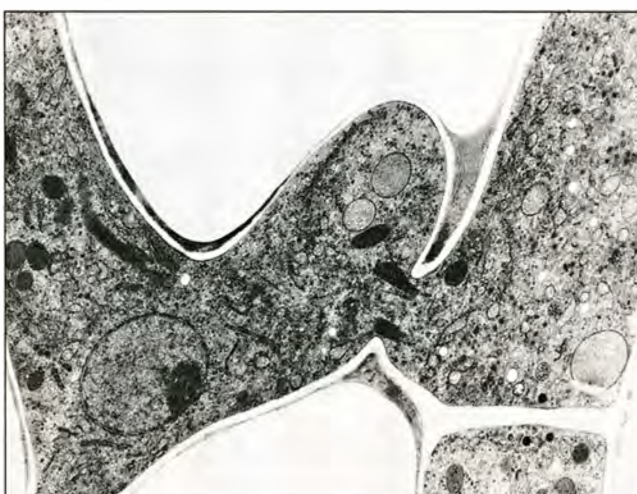
Luego, se obtuvieron partículas similares de otras líneas europeas hipovirulentas. Sin embargo, no se ha logrado extraer ninguna de líneas virulentas, ya fuesen europeas o americanas, ni de cepas hipovirulentas norteamericanas. Y sin embargo, como veremos a continuación, hay pruebas de que las partículas VLP están siempre presentes allí donde se da hipovirulencia.

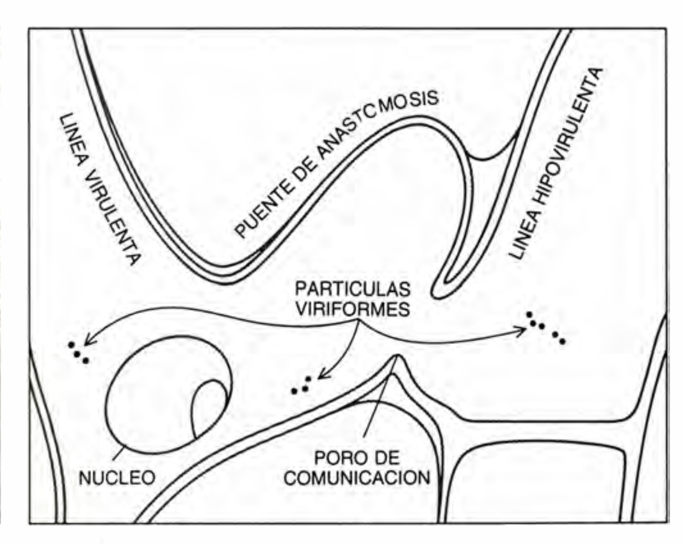
No cabía duda de que podría aprenderse mucho sobre las partículas VLP estudiando su posición dentro del citoplasma fúngico, pero durante muchos años resultaron vanos los esfuerzos realizados para observarlas in situ, debido a que los procesos de fijación química habituales no respetaban las estructuras subcelulares delicadas. En 1981 se me ocurrió que podrían obtenerse instantáneas microscópicas de partes del hongo gracias a la técnica de sustitución por congelación. Esta técnica permite fijar las estructuras celulares de forma permanente, realizando primero una congelación instantánea y sumergiéndolas luego en productos químicos que deshidratan y estabilizan el citoplasma y su contenido. A continuación se embebe el tejido en plástico, se secciona, se tiñe y se observa con un microscopio electrónico de transmisión.

Procediendo así, encontré partículas VLP en las hifas y conidios (esporas asexuales) de líneas hipovirulentas europeas y norteamericanas, pero no en las de líneas virulentas, idénticas genéticamente, que estaban libres de ARNdc. Las partículas eran esféricas y estaban delimitadas por una membrana; aunque menores y de morfología distinta de las que Dodds había aislado, creo que muy probablemente son las mismas. Para empezar, el microscopio electrónico de transmisión siempre hace que las partículas cuando están aisladas parezcan mayores que cuando pertenecen a una sección de tejido. En segundo lugar, pruebas citoquímicas pusieron de manifiesto que las partículas contienen ácido nucleico y que están rodeadas por una membrana lipídica, membrana muy susceptible a la deformación. Por tanto, el procedimiento de extracción de Dodds podría haber alterado el aspecto de las partículas.

En las líneas hipovirulentas europeas, las partículas VLP formaban agregados rodeados por retículo endoplasmático rugoso y otros elementos de la maquinaria celular de síntesis. En las líneas norteamericanas, las partículas VLP eran escasas y estaban dispersas por el citoplasma, lo que puede explicar los fracasos anteriores al intentar extraerlas de las hifas.

Las partículas VLP sólo se parecen a un componente celular de los que se encuentran normalmente en *C. parasitica*, a saber, las vesículas apicales, que ayudan a producir membranas y paredes celulares en los extremos en crecimiento de las hifas. Esta semejanza me lleva a sospechar que el mecanismo de crecimiento ha sido desviado aquí para que cumpla otro pro-





3. ANASTOMOSIS de dos cepas fúngicas cuyas estructuras fueron congeladas y fijadas químicamente, cuatro horas después de su inicio. Se trata de una microfotografía obtenida con el microscopio electrónico de transmisión. Las hifas de cada línea, virulenta la una e hipovirulenta la otra, han formado un puente a través del cual pueden intercambiar fluido celular (*dibujo*). Las partículas viriformes (*marcadas con flechas*) están presentes en el puente y en la hifa de la línea virulenta, indicando que el ARN de doble cadena (ARNdc) ya se ha transferido a dicha cepa virulenta.

pósito: el de proteger las células del hongo contra el ácido nucleico extraño, el ARNdc.

Desde este punto de vista, las partículas VLP son productos finales, formados cuando se encierra un ácido nucleico extraño en estructuras hechas con precursores de la pared celular. En la fase final, las partículas VLP podrían emigrar a la membrana celular y fusionarse con ella, depositando así su contenido en la pared celular, con lo que se purgaría el ARN bicatenario de la célula.

Pero, si realmente el ARNdc de las partículas VLP está neutralizado, ¿qué es lo que causa las aberraciones morfológicas de las líneas de *C. parasitica* infectadas con ARNdc? Quizás la mayor parte del ARNdc de las líneas hipovirulentas nunca queda atrapado en las partículas VLP, sino que flota libremente en el citoplasma, como lo hacen los viroides. Van Alfen sugiere que el ARN bicatenario puede representar el genoma de un virus deficiente, virus que ha perdido la capacidad de producir una cápside o cubierta proteica. En cualquier caso, el ARNdc asociado con la hipovirulencia es único entre los agentes biológicos que se sabe que infectan hongos.

La técnica de congelación también me permitió estudiar la anastomosis hifal entre varias líneas de *C. parasitica*. Pude confirmar, por primera vez, que existe continuidad citoplasmática entre líneas compatibles vegetativamente. Además, encontré partículas VLP en el puente de anastomosis y en la línea virulenta sólo cuatro horas después de que comenzase la anastomosis [*veáse la figura 3*]. A pesar de que las líneas incompatibles también inician la anastomosis, su citoplasma degenera pronto tras la fusión, lo que impide el paso del ARNdc y, por tanto, del carácter de hipovirulencia de una raza a otra.

Las herramientas de la biología molecular han permitido a los investigadores descubrir algunos secretos de la hipovirulencia. Un grupo dirigido por Robert E. Rhoads, de la Universidad de Kentucky, ha usado una técnica llamada hibridación, en un intento de unir ARNdc de una cepa hipovirulenta, marcado radiactivamente, a ARNdc de otras cepas hipovirulentas. Una hibridación con éxito indica que la sonda y su blanco guardan un estrecho parentesco. El grupo de Rhoads encontró, sin embargo, que la hibridación se produce únicamente cuando la sonda y su blanco proceden de líneas originarias del mismo continente, ya sea Europa o Norteamérica. Este resultado demuestra que la hipovirulencia surgió independientemente en los dos continentes y que implica dos grupos distintos de ARN bicatenario.




4. CHANCRO EN EXPANSION causado por una línea virulenta de *Cryphonectria parasitica*. El castaño acabará muriendo (*izquierda*). Si se inocula el chancro con líneas hipovirulentas, el hongo virulento se debilita, permitiendo que el árbol aísle la infección con tejido de callo (*fotografía de la derecha*).

Estudios de hibridación adicionales encontraron dos líneas hipovirulentas en Michigan cuyos ARNdc no se unían con el ARNdc de otras muchas razas del país. Puede haber, por tanto, muchos tipos no emparentados de ARNdc en líneas de *C. parasitica* hipovirulentas, al menos en Estados Unidos. Sin embargo, otros estudios de hibridación han puesto de manifiesto que, en algunas líneas hipovirulentas, los segmentos múltiples de ARNdc están emparentados entre sí. Este hallazgo llevó a Donald L. Nuss, del Instituto Roche de Biología Molecular de Nutley, N. J., a conjeturar que el ARN bicatenario pudiera ser menos complejo genéticamente de lo que parece. Si así fuera, se facilitaría la identificación de los genes responsables de la hipovirulencia.

Para acotar con precisión el segmento, o la combinación de segmentos, de ARNdc que origina la hipovirulencia, se debe determinar primero cómo afecta el ARNdc a la célula. Algunos biólogos moleculares del Instituto Roche han empezado a estudiar este problema por secuenciación de una cadena de ARNdc de una cepa hipovirulenta europea. Tradujeron in vitro una parte de la cadena, produciendo una proteína con un peso molecular de 29 kilodalton. Esta proteína fue el primer producto sintetizado a partir del ARNdc asociado con la hipovirulencia.

Para determinar la forma en que los productos proteicos del ARNdc inducen la hipovirulencia en *C. parasitica*, los investigadores usan ahora una enzima, llamada transcriptasa inversa; obtienen así una copia de ADN de los genes del ARNdc que especifican la producción de proteínas individuales. Luego insertan este ADN en cepas virulentas, proceso llamado transformación, con lo que intentan descubrir el efecto que cada proteína tiene sobre el hongo.

Pertrechados con esta información, los fitopatólogos podrán determinar si el mecanismo causante de la hipovirulencia condena necesariamente al hongo a una debilidad reproductiva. Si no es así, la ingeniería genética podría ser capaz de crear líneas lo suficientemente débiles como para no matar los castaños, pero lo suficientemente fuertes para competir con las líneas virulentas del bosque. Tal vez puedan descifrarse los determinantes genéticos de la incompatibilidad vegetativa, de modo que la hipovirulencia pueda transmitirse a todas las líneas de *C. parasitica*. Ello podría conseguirse mediante el estudio de unas líneas hipovirulentas de reciente descubrimiento, llamadas multiconvertidoras, que tienen capacidad para pasar la hipovirulencia a los miembros de diversos grupos vegetativamente incompatibles.

Los biólogos disponen de otro camino para intentar salvar al castaño americano: pueden tratar de producir un árbol más resistente, aunque sigan trabajando para debilitar su parásito. Algunos expertos en genética vegetal están cruzando el castaño con sus parientes asiáticos más resistentes, vol-

viendo a cruzar luego los híbridos resultantes con reserva pura americana, en un esfuerzo por conseguir las cualidades deseadas: la belleza del árbol americano y la resistencia a la enfermedad del árbol asiático. Otros investigadores han realizado innumerables búsquedas para encontrar castaños que hayan sobrevivido al chancro, sobre todo en Carolina del Norte y en Virginia. Se injertan esquejes tomados de los supervivientes en plantores de individuos resistentes al chancro. A pesar de todo, tal cultivo es lento y sus resultados inciertos.

Es un campo en el que la biología molecular también puede ser de ayuda algún día. Cuando se identifiquen los genes que confieren resistencia, debería ser posible incorporarlos a plantones de los que terminarían naciendo castaños americanos resistentes. Ni la mejora ni la introducción de genes pueden devolver al castaño su gloria inicial de la noche a la mañana, pues otras especies han llenado el vacío ecológico originado por la destrucción del árbol. Sin embargo, su reintroducción en los jardines de aquellos que lo aman sería una simpática victoria. Entonces se cumpliría la profecía de Robert Frost:

¿Acabará el chancro con el castaño?
Los agricultores se inclinan a creer que no.
Sigue palpitando en sus raíces
y emitiendo nuevos brotes,
hasta que aparezca otro parásito
que acabe con la enfermedad.

BIBLIOGRAFIA COMPLEMENTARIA

DEATH OF THE AMERICAN CHESTNUT. George H. Hepting en *Journal of Forest History*, vol. 18, n.° 3, págs. 60-67; julio de 1974.

PROCEEDINGS OF THE AMERICAN CHESTNUT SYMPOSIUM. Universidad de Virginia occidental, Morgantown, 4-5 de enero de 1978. Dirigido por William L. MacDonald et al. Universidad de Virginia occidental, 1978.

CHESTNUT BLIGHT, OTHER ENDOTHIA DISEASE, AND THE GENUS *ENDOTHIA*. Martha K. Roane, Gary J. Griffin y John R. Elkins. Series monográficas de la Sociedad Fitopatológica americana. St. Paul, Minn., APS Press, 1986.

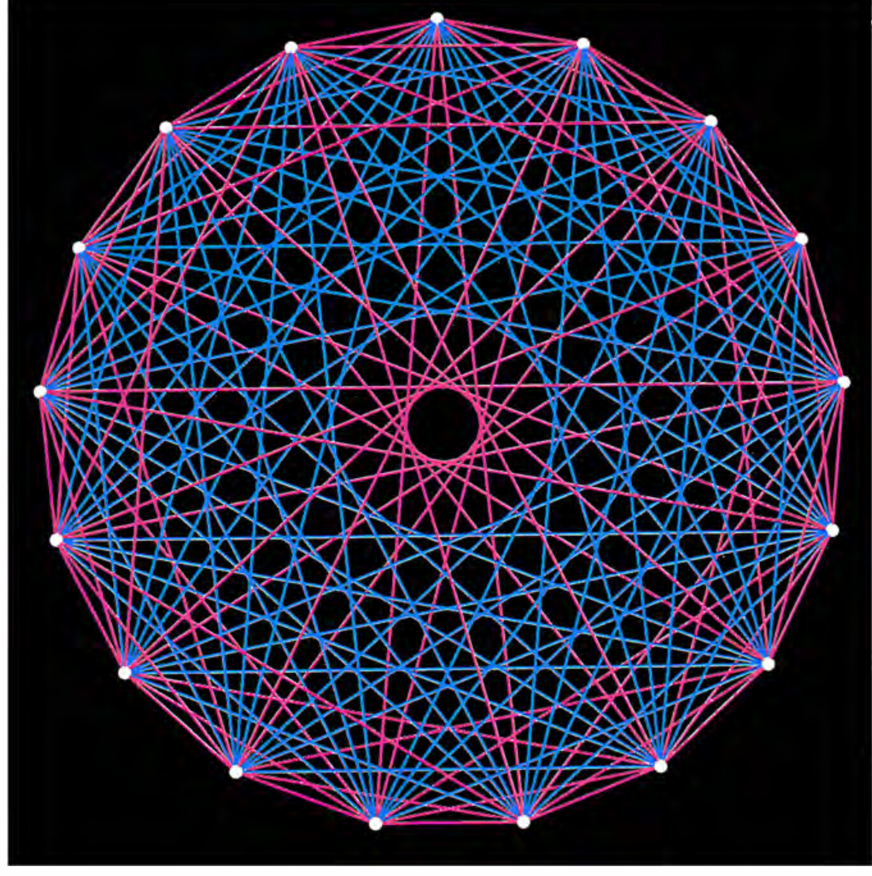
VIRUS-LIKE PARTICLES IN HYPHAE AND CONIDIA OF EUROPEAN HYPOVIRULENT (DSRNA-CONTAINING) STRAINS OF *CRYPHONECTRIA PARASITICA*. Joseph R. Newhouse, William L. MacDonald y Harvey C. Hoch en *Canadian Journal of Botany*, vol. 68, n.° 1, págs. 90-101; enero de 1990.

Teoría de Ramsey

Frank Plumpton Ramsey, un brillante matemático, demostró que era imposible el desorden completo. Toda colección grande de números, puntos u objetos contiene, por necesidad, configuraciones regulares

Ronald L. Graham y Joel H. Spencer

Según una tablilla cuneiforme de hace 3500 años, un antiguo sabio sumerio alzó la mirada a las estrellas del cielo y vio un león, un toro y un escorpión. Es probable que un astrónomo moderno describiera las constelaciones diciendo que se trata de agrupaciones transitorias de estrellas, que nosotros los terrícolas observamos desde el confín de una galaxia bastante vulgar. No obstante, la mayoría de los amantes de la observación celeste estarían de acuerdo en que el cielo nocturno parece lleno de constelaciones en forma de líneas rectas, rectángulos y pentágonos. ¿No podría suceder que tales configuraciones geométricas fueran fruto de fuerzas desconocidas del cosmos?



1. PROBLEMA DE LA REUNION, típico de la teoría de Ramsey. ¿A cuántas personas hay que invitar para que se forme con certeza un grupo de cuatro que se conocen dos a dos, o un grupo de cuatro que no se conocen mutuamente? En el diagrama, las personas se representan mediante puntos. Un lado rojo conecta a dos personas que ya se conocen, mientras que un lado azul enlaza a cada dos que no se conocen. En el grupo de 17 personas mostrado arriba no hay cuatro puntos cuya red de lados sea exclusivamente roja o exclusivamente azul. Por tanto, se necesitan más de 17 personas para garantizar que habrá cuatro que se conozcan o cuatro que no. De hecho, en cualquier grupo de 18 personas hay siempre cuatro que se conocen mutuamente o cuatro que se desconocen dos a dos.

Las matemáticas proporcionan una explicación mucho más verosímil. En 1928, Frank Plumpton Ramsey, matemático, filósofo y economista británico, demostró que en realidad tales configuraciones se hallan implícitas en toda configuración grande, ya se trate de un grupo de estrellas, una formación de guijarros o una serie de números obtenidos al azar por lanzamiento de un dado. Ante un número suficientemente grande de estrellas, por ejemplo, podremos siempre hallar un grupo que venga a dibujar una configuración particular: una línea recta, un rectángulo o, puestos a eso, una Osa Mayor. La teoría de Ramsey afirma, en efecto, que cualquier estructura contendrá necesariamente una subestructura organizada. Como ya proclamara hace unos 25 años el matemático estadounidense Theodore S. Motzkin, la teoría de Ramsey comporta que el desorden completo sea una imposibilidad.

Los especialistas en teoría de Ramsey se esfuerzan por determinar cuál es el número preciso de estrellas, números o cifras necesarios para garantizar la existencia de cierta subestructura deseada. La solución de tales problemas suele requerir decenios, y tan sólo ceden ante los razonamientos más delicados e ingeniosos. Al tiempo que buscan soluciones, los teóricos ayudan a los ingenieros en sus esfuerzos por construir mejores redes de comunicaciones, amén de sistemas de transmisión y recuperación de información. Los teóricos han descubierto también algunas de las herramientas matemáticas que servirán de guía a los científicos en el siglo venidero. Y lo que quizás sea más importante, la teoría de Ramsey está sondeando la

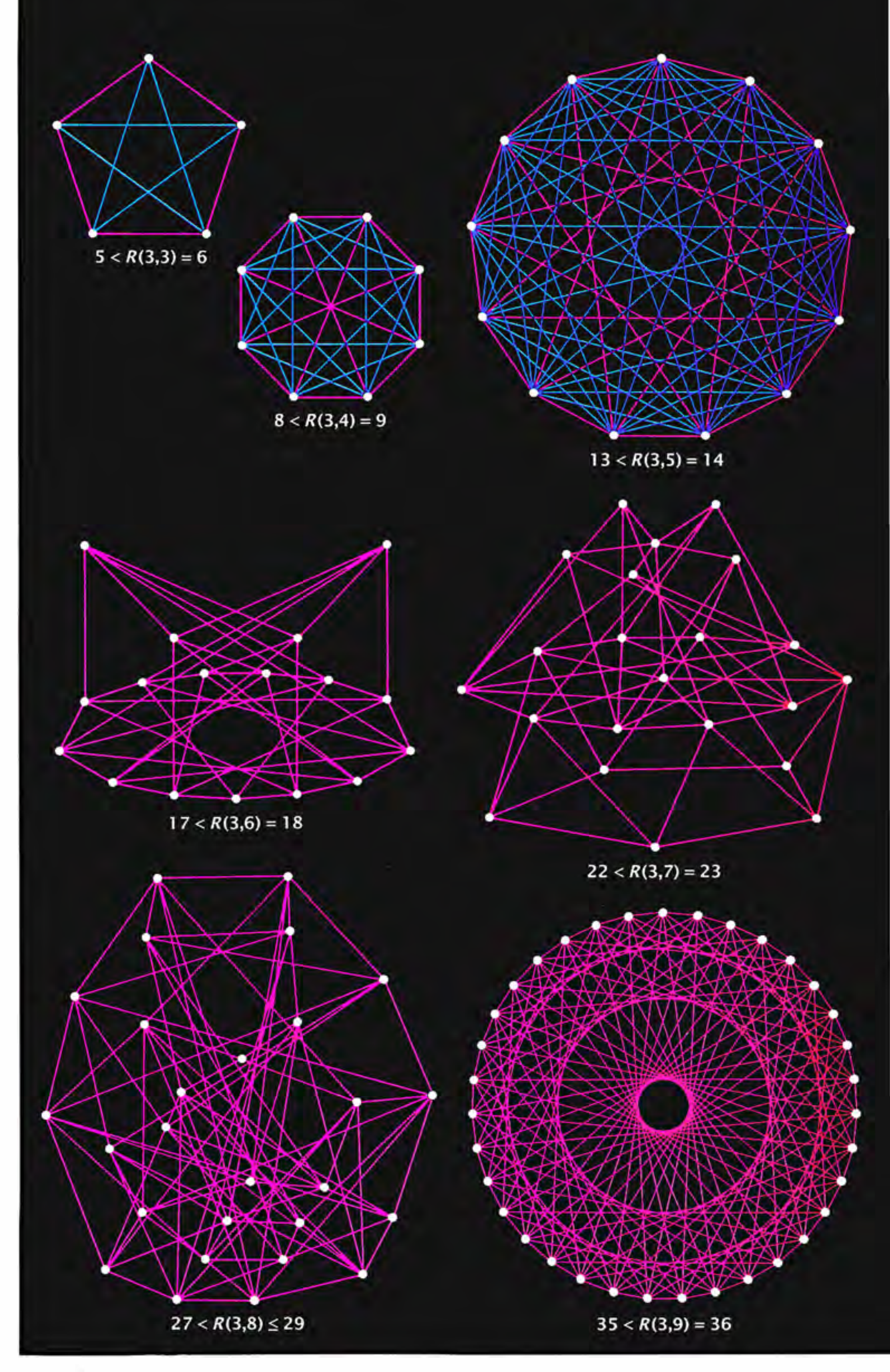
RONALD L. GRAHAM y JOEL H. SPENCER han escrito, con Bruce L. Rothschild, de la Universidad de California en Los Angeles, el tratado definitivo sobre teoría de Ramsey. Graham, director de investigación de la división de ciencias informáticas en los Laboratorios AT & T Bell, enseña matemática en la Universidad de Rutgers. Se recibió de doctor por la Universidad de California en Berkeley en 1962. Spencer es profesor de matemáticas e informática en el Instituto Courant de Ciencias Matemáticas de la Universidad de Nueva York. Obtuvo su doctorado por la Universidad de Harvard en 1970. Muestra particular interés por la historia de la matemática húngara, como la obra de Paul Erdös, que tantos nuevos campos abrió en teoría de Ramsey.

estructura última de las matemáticas, una estructura que trasciende a la del universo.

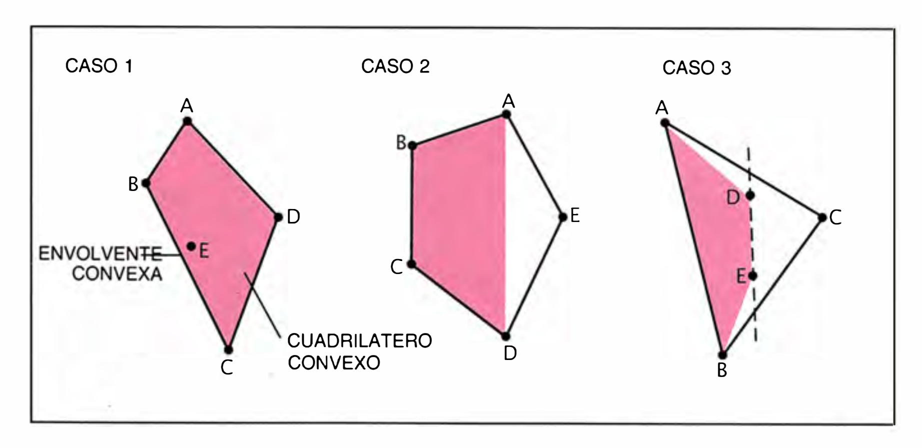
A diferencia de muchas ramas de la matemática en que hoy se ocupan los profesionales, la teoría de Ramsey se presta a una presentación intuitiva. Más aún, el encanto de la teoría de Ramsey emana, en parte, de la simplicidad con que se pueden formular los problemas. Por ejemplo, si se eligen al azar seis personas (sean Alfredo, Beni, Carlos, Diana, Eduardo y Fernanda), ¿es seguro que tres de ellas se conocerán entre sí o que habrá tres de ellas que son mutuamente desconocidas?

Podemos resolver de muchos modos el "problema de la reunión". Podríamos, por ejemplo, enumerar exhaustivamente todas las combinaciones concebibles y examinarlas una por una, para ver si contienen un grupo de tres conocidos o de tres desconocidos. Pero, dado que tendríamos que examinar 32.768 (= 2^{15}) combinaciones, este método de fuerza bruta ni es práctico ni proporciona un buen entendimiento de la situación.

Por fortuna para nosotros, podemos hallar la solución considerando dos casos sencillos. Supongamos, en el primero, que Alfredo conoce al menos a tres de los demás, por ejemplo, a Beni, Carlos y Diana. Si Beni y Carlos, o Beni y Diana, o Carlos y Diana, se conocen, resultará que Alfredo más el par de conocidos forman un grupo de tres personas que se conocen entre sí. En caso contrario, Beni, Carlos y Diana son mutuamente desconocidos. En el segundo caso, supongamos que Alfredo sólo conoce como máximo a dos de los demás, por ejemplo, a Beni y a Carlos. Si Diana



2. LOS NUMEROS DE RAMSEY son, por definición, el mínimo valor de n para que en un grupo de n puntos haya con certeza un grupo de j puntos que forme una red completa de lados rojos o un grupo de k puntos que constituya una red completa de lados azules. Los diagramas de la figura indican la magnitud de tales números. El primer diagrama muestra cinco puntos conectados por lados rojos y azules, de modo tal que ningún grupo de tres forma una red completa de color rojo o de color azul. El primer diagrama implica, pues, que el número de Ramsey para tres rojo y tres azul tiene que ser mayor que cinco. De forma similar podríamos aducir que el segundo diagrama sugiere que el número de Ramsey para tres rojo y cuatro azul es mayor que ocho. Merced a otras complejas técnicas, puede demostrarse que el número de Ramsey correspondiente a tres rojo y tres azul es seis, en tanto que el número correspondiente a tres rojo y cuatro azul es nueve. Se dan en la figura todos los números de Ramsey que son exactamente conocidos, menos el número de Ramsey para cuatro rojo y cuatro azul, cuyo diagrama vemos ilustrado en la figura 1. (Se han omitido los lados azules en algunos de estos diagramas por mor de mayor claridad.) Se ha demostrado que el número de Ramsey para tres rojo y ocho azul es mayor que 27 y menor o igual que 29. Se mostró recientemente (aunque no se ha verificado todavía) que tal número es 28. (Recreación de Gabor Kiss.)



3. REDESCUBRIMIENTO DE LA TEORIA DE RAMSEY en 1933, cuando Esther Klein, una joven estudiante, introdujo un nuevo problema geométrico: "Si en un plano yacen cinco puntos sin que nunca tres estén alineados, demostrar que cuatro de ellos formarán siempre un cuadrilátero convexo". Todos los casos del problema son variantes de los tres aquí mostrados. El más sencillo se da cuando la envolvente convexa —el polígono convexo que encierra todos los puntos— es un cuadrilátero. Si la envolvente convexa es un pentágono, se podrán conectar cuatro puntos cualesquiera para formar un cuadrilátero. Si la envolvente convexa es triangular, habrá de contener a dos de los puntos, en este caso, *D* y *E*. La línea *DE* escinde el triángulo de modo tal que dos puntos, *A* y *B*, se encuentran en un mismo semiplano. Los cuatro puntos *ABDE* formarán un cuadrilátero convexo.

y Eduardo, o Diana y Fernanda, o Eduardo y Fernanda no se conocen, Alfredo más la pareja de no-conocidos forman un grupo de tres personas en el que nadie conoce a los demás. De no ser así, Diana, Eduardo y Fernanda se conocen mutuamente. Con sólo seis frases hemos demostrado por qué, en todo grupo de seis personas, es seguro que habrá al menos tres que se conozcan o al menos tres que no se conozcan. Y lo que va más al caso, la solución del problema de la reunión constituye un caso particular de teoría de Ramsey.

Podemos enunciar el teorema completo por generalización de este caso particular. En lugar de considerar seis personas en el problema, podemos tener un número cualquiera de ellas o, si se quiere, cualquier número de objetos. No hay necesidad de limitarse a dos relaciones, la de conocerse o la de no conocerse; podemos tener cualquier número de relaciones mutuamente exclusivas, verbigracia, amigos, enemigos y neutrales.

Podemos describir ahora el teorema de Ramsey en toda su plenitud. Si el número de objetos de un conjunto es suficientemente grande y cada par de objetos posee sólo una de cierto número de relaciones, habrá siempre un subconjunto compuesto por cierto número de objetos donde cada par presente la misma relación.

Frank Ramsey, el primero en demostrar la proposición anterior, en 1928, se crió en Cambridge, Inglaterra. Su padre, Arthur S. Ramsey, era profesor de matemáticas y presidente del Colegio de la Magdalena de Cambridge. El joven Ramsey fue en 1925 el más distinguido alumno de matemáticas que se graduó en la universidad. Aunque se entregó sobre todo a la filosofía y la lógica matemática, efectuó también contribuciones en economía, probabilidad, teoría de la decisión, psicología cognitiva y semántica.

Al poco de su graduación ingresó en un grupo de economistas encabezado por John Maynard Keynes. Ramsey escribió tan sólo dos artículos sobre economía matemática, pero ambos son todavía hoy ampliamente citados. En filosofía, su inspiración bebe, sobre todo, en George E. Moore, Ludwig Wittgenstein y Bertrand Russell. Moore escribió: "Era un pensador extraordinariamente claro: nadie supo eludir mejor que él la clase de confusiones mentales en las que incurren hasta los mejores filósofos". Mas, trágicamente, en 1930, contando sólo 26 años, Ramsey cayó enfermo y murió de las complicaciones postoperatorias de una intervención quirúrgica abdominal.

La historia de cómo dedujo Ramsey su teoría, un par de años antes de su muerte, tiene una pincelada irónica. Ramsey dio con la idea central mientras trataba de demostrar una premisa planteada por Russell y Alfred North Whitehead en su obra maestra, *Principia Mathematica*. Proponían en ella que todas las verdades matemáticas podían deducirse de un sistema restringido de axiomas. David Hilbert llevó más allá esta idea y sugirió que tendría que existir un procedimiento para decidir si una proposición dada era o no consecuencia de un conjunto de axiomas. Ramsey demostró que tal procedimiento sí existía para un caso particular. (Pocos años después, Kurt Gödel, seguido por el matemático inglés Alan M. Turing y otros, demostró concluyentemente que no existía tal procedimiento de decisión en el caso general.)

Ramsey demostró su teorema como un primer paso en su intento de probar el caso particular. Según luego se vio, pudo haber conseguido el mismo objetivo por otros medios. Ramsey había demostrado un teorema que era superfluo para un argumento; argumento que nunca hubiera podido demostrar en el caso general.

Así quedaron las cosas hasta 1933, cuando dos matemáticos húngaros, Paul Erdös y George Szekeres, redescubrieron la teoría de Ramsey. Fueron ellos, en gran medida, los responsables de su popularización en la comunidad matemática. Erdös tenía a la sazón 19 años, y era alumno de la Universidad de Budapest; Szekeres se había graduado hacía poco en ingeniería química por el Politécnico de la capital húngara. Ellos dos se reunían, con un grupo de compañeros de estudios, casi todos los domingos en un parque o en la escuela para hablar, sobre todo, de matemáticas.

En una reunión del invierno de 1933, una de las estudiantes, Esther Klein, retó al grupo a resolver un curioso problema: si se dan cinco puntos del plano situados de manera que nunca haya tres alineados, demostrar que cuatro de estos puntos formarán siempre un cuadrilátero convexo. (El término convexo sugiere figuras geométricas sin entrantes, como un hexágono; una estrella de cinco puntas, por el contrario, no es convexa. De manera más precisa, un polígono es convexo si cada segmento rectilíneo trazado entre dos cualesquiera de sus puntos yace dentro del polígono.)

Tras dejar que sus amigos reflexionasen en el problema, Klein les expuso una demostración [*véase la figura 3*]. Erdös y Klein obtuvieron una pronta generalización del problema. Se dieron cuenta de que cinco de cada nueve puntos del plano formarían siempre un pentágono convexo. Propusieron entonces un nuevo problema: si el número de puntos yacentes en el plano es igual a $1 + 2^{k-2}$, donde k es igual a 3, 4, 5, y así sucesivamente, ¿se podrán seleccionar siempre k puntos que dibujen un polígono convexo de k lados?

Szekeres rememoraba la escena en una monografía. "Pronto nos percatamos de que un razonamiento sim-

plista no serviría de nada; sentíamos la excitación de haber descubierto en nuestro círculo un problema geométrico de nuevo tipo." Szekeres, ardorosamente, hizo ver que siempre existe un número n tal que, si n puntos yacen en el plano de modo que nunca haya tres en línea recta, es posible seleccionar entre ellos k puntos que formen un polígono convexo de k lados. Dicho de otro modo, dados suficientes puntos, siempre podremos hallar un conjunto que forme un polígono particular. Al demostrar esta proposición, Szekeres había redescubierto el teorema de Ramsey, aunque ninguno de los miembros de su grupo lo supiera en aquel entonces.

Erdös y Szekeres dieron a conocer sus resultados en 1934. Pero, hasta ahora, ni ellos ni nadie han podido demostrar la conjetura de Erdös según la cual $n = 1 + 2^{k-2}$ son puntos suficientes. Erdös suele referirse a su publicación conjunta con el remoquete del "artículo de final feliz", porque, al poco de ser publicado, Klein y Szekeres se casaron. Erdös se convirtió en el más prolífico matemático de este siglo.

A Erdös le tenía intrigado la idea de Ramsey de que cualquier estructura suficientemente grande debía contener una subestructura regular de tamaño dado. Se preguntó, no obstante, cuán grande exactamente habría de ser la estructura ambiente para garantizar cierta subestructura. En consecuencia, Erdös empezó a trabajar en una versión del problema de la reunión.

En esta versión, las seis personas están representadas mediante seis puntos. Por conveniencia, los puntos se dibujan en el plano de manera que nunca haya tres alineados. Los puntos están conectados por un lado, que se colorea para representar la relación de las dos personas correspondientes. Un lado rojo significa que las dos personas se conocen, mientras que un lado azul significa que no se conocen.

Por tanto, si tres personas se conocen entre sí, los lados que conectan los puntos correspondientes formarán un triángulo rojo, mientras que, si tres personas no conocen a ninguna de las otras dos, definirán un triángulo azul. El problema de la reunión admite la siguiente reformulación: ¿Es cierto que, si cada uno de los lados que conectan seis puntos se colorea arbitrariamente de rojo o de azul, se formará siempre un triángulo rojo o un triángulo azul?

El problema que Erdös estudió es una versión general de éste. Erdös define una red completa como una colección finita de puntos en la que todos están conectados con todos mediante lados. Preguntó entonces cuál sería la mínima red completa que, al ser arbitrariamente coloreada, garantizaría una red completa (roja o azul) de tres puntos. La solución es la red completa de seis vértices. El problema y su solución se expresan así con mayor propiedad: el número de Ramsey correspondiente a tres rojo y tres azul es igual a seis.

¿Qué se sabe del número de Ramsey correspondiente a cinco rojo y tres azul? O con otras palabras, ¿cuál es la mínima red completa que, arbitrariamente coloreada de rojo y de azul, garantizará una red roja de cinco puntos o una red azul de tres puntos? El número de Ramsey para cinco rojo y tres azul es 14, hecho no demostrado hasta 1955 por Robert E. Greenwood, de la Universidad de Texas en Austin, y Andrew M. Gleason, de la Universidad de Harvard.

Los números de Ramsey son notoriamente difíciles de calcular. Los esfuerzos de generaciones de matemáticos y ordenadores solamente han tenido éxito en hallar siete números de Ramsey [*figura 2*].

Para expresar la dificultad de calcular números de Ramsey, Erdös suele imaginar la siguiente situación. Los extraterrestres invaden la Tierra y amenazan con aniquilarla dentro de un año, a menos que los humanos logren hallar el número de Ramsey correspondiente a cinco rojo y cinco azul. Podríamos reclutar las mentes mejor dotadas del mundo y los ordenadores más rápidos; quizás así lográsemos calcular el número de Ramsey en el plazo de un año. Empero, si los extraterrestres exigieran el número de Ramsey para rojo seis y azul seis, no tendríamos más opción que lanzar un ataque preventivo.

No obstante, Erdös sí halló una forma para hacerse una idea de cuán grande ha de ser un número de Ramsey. ¿Qué sucedería si pudiese hallar un coloreado rojo y azul de una gran red completa que *no* formase ni una red roja ni una red azul de tres puntos? Vemos un tal coloreado de la red completa de cinco vértices en la figura 2. Se sigue que el número de Ramsey correspondiente a rojo tres y azul tres debe ser mayor que 5. El 5 es cota inferior del número de Ramsey.

En 1947, Erdös propuso un método

La teoría de Ramsey y las progresiones aritméticas

Una progresión aritmética es una sucesión de números en la cual la diferencia de dos términos consecutivos permanece constante. Por ejemplo, 7, 10, 13, 16 es una progresión aritmética cuya diferencia entre términos consecutivos es tres. De la teoría de Ramsey se sigue esta proposición sobre progresiones aritméticas: si en la secuencia natural de 1 a 9 cada uno de los números es coloreado de rojo o de azul, habrá al menos tres números rojos o tres números azules que formen progresión aritmética.

Para demostrar la conjetura podríamos verificarla en cada una de las 512 formas de colorear los nueve números en dos tonos. Pero también es posible demostrarla considerando sólo dos casos. Empezamos con el caso de que 4 y 6 sean del mismo color, azul por ejemplo.

1 2 3 4 5 6 7 8 9

Para impedir que se forme la progresión aritmética 4, 5, 6, coloreamos el 5 de rojo.

1 2 3 4 5 6 7 8 9

Para impedir las progresiones 2, 4, 6, y 4, 6, 8, coloreamos de rojo al 2 y al 8.

1 2 3 4 5 6 7 8 9

lo cual deja la progresión aritmética roja 2, 5, 8. Así, pues, si 4 y 6 son del mismo color, existirá siempre una progresión aritmética roja o una azul. Consideraremos seguidamente el caso de que 4 y 6 sean de distinto color. Podemos colorear al 5 de rojo o de azul sin formar progresión aritmética, por lo que arbitrariamente optamos por colorear el 5 de rojo.

1 2 3 4 5 6 7 8 9

Proseguimos el coloreado de los números como sigue

3 para evitar **3 4 5**	**9** para evitar **3 6 9**
7 para evitar **5 7 9**	**8** para evitar **6 7 8**
2 para evitar **2 5 8**	**1** para evitar **1 2 3**

Este coloreado da la secuencia

1 2 3 4 5 6 7 8 9

Empero, nos queda todavía la progresión 1, 5, 9. Así, pues, independientemente de si 4 y 6 son o no del mismo o de distinto color, existe siempre una progresión aritmética azul o roja.

insólito de hallar una cota inferior de cualquier número de Ramsey: lanzar una moneda. Realizó un experimento mental, en el que cada uno de los lados de una red completa de un millón de vértices, pongamos por caso, era coloreado según el resultado de lanzar una moneda insesgada. El lado sería rojo de salir "cruz", y azul de salir "cara". Seguidamente, trató de demostrar que el número de Ramsey correspondiente, sea por caso, a 34 rojo y 34 azul es mayor que un millón. El experimento sería un éxito si no apareciera ni una red roja ni una red azul de 34 puntos.

¿Cómo podría asegurarse el éxito? Dados 34 puntos, tiene que haber 561 lados que los conecten. Si el primer lanzamiento de la moneda especifica que el primer lado ha de ser azul, para producir una red azul también los 560 lanzamientos siguientes han de especificar azul. La probabilidad de que así ocurra es de 1/2 elevado a 561. La probabilidad de que aparezca una red roja es la misma, por lo que la probabilidad total es el doble, o sea, aproximadamente igual a $2{,}6 \times 10^{-169}$.

Ahora bien, el número de conjuntos de 34 puntos que pueden extraerse de un total de un millón de puntos es igual a $(1.000.000 \times 999.999 \times \ldots \times 999.967)$ dividido entre $(34 \times 33 \times \ldots \times 2 \times 1)$, que es alrededor de $3{,}4 \times 10^{165}$. Así, pues, de todas las posibles redes completas de 34 puntos, sería de esperar que unas $(3{,}4 \times 10^{165}) \times (2{,}6 \times 10^{-169})$ —o sea, alrededor de 0,001— fuesen monocromáticas. Por tanto, el 99,9 por ciento de las veces el experimento conceptual constituiría un éxito, al no crearse conjuntos monocromáticos de 34 puntos.

Erdös aplicó luego una sutil reducción al absurdo. Tomó como hipótesis que ningún sistema de coloreado tendría éxito. En tal caso, el experimento conceptual habría de tener probabilidad de éxito igual a 0, lo que se sabe que no es así. La hipótesis tiene que ser incorrecta; ha de existir un coloreado con éxito (no con mera probabilidad del 99,9 por ciento, sino con certeza absoluta). La existencia del coloreado implica que un millón es una cota inferior para el número de Ramsey correspondiente a 34 rojo y 34 azul.

Este razonamiento, conocido por "método probabilístico", ha proporcionado las mejores cotas inferiores conocidas para los números de Ramsey. Empero, el método probabilístico no da indicación alguna acerca de la "construcción" efectiva del coloreado deseado. Los investigadores, en sus esfuerzos por obtener tales construcciones, han recurrido a toda una gama de técnicas, que van desde la teoría de números, pasando por la teoría de conjuntos y otras ramas de las matemáticas. Los resultados, prometedores, no se aproximan todavía a las cotas obtenidas por lanzamiento de la moneda.

Aunque gran parte de los primeros trabajos en teoría de Ramsey se concentraron en conjuntos de puntos y rectas, muchos problemas se referían a conjuntos de números. Así, el matemático holandés Bartel L. van der Waerden comenzó a resolver tales problemas antes incluso de que Ramsey demostrase su teorema.

Van de Waerden tuvo noticia, en 1926, de un curioso problema relativo a progresiones aritméticas. Como su nombre da a entender, una progresión geométrica es una sucesión de números en la que la diferencia entre términos consecutivos se mantiene constante. Por ejemplo, la sucesión 3, 5, 7 es una progresión aritmética de tres términos en la cual la diferencia entre términos consecutivos es 2. El caso particular del problema que captó el interés de van der Waerden fue el siguiente: Si cada uno de los dígitos de 1 a 9 se imprime en una página en uno de dos colores, por ejemplo, rojo o azul, ¿será siempre cierto que habrá tres números rojos o tres números azules en progresión aritmética? Podemos ver la solución en el recuadro de la página anterior.

Van der Waerden se desafió a sí mismo con la siguiente generalización: si n es un entero suficientemente grande, y si cada uno de los enteros de 1 a n se imprime al azar de uno de dos colores, habrá siempre una sucesión aritmética monocromática con cierto número de términos. Se podría pensar que esta proposición no es más que el teorema de Ramsey para progresiones aritméticas, aunque por lo general se le conoce como teorema de van der Waerden.

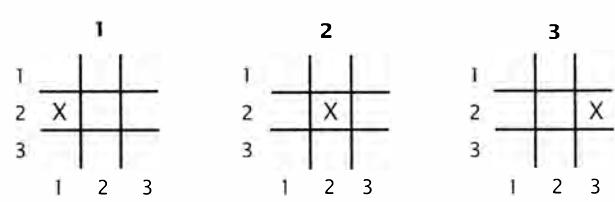
Van der Waerden reclutó la ayuda de sus colegas Emil Artin y Otto Schreier. Más tarde escribiría: "Fuimos al despacho de Artin en la facultad de matemáticas de la Universidad de Hamburgo, y tratamos de hallar una demostración. Trazamos algunos diagramas en la pizarra. Teníamos lo que los alemanes llaman 'Einfälle': ideas repentinas surgidas como destellos en nuestras mentes. Ideas nuevas que dieron varias veces nuevos giros a la discusión, y una de las cuales condujo finalmente a la solución."

La teoría de Ramsey y el tres en raya

En 1926, Bartel van der Waerden demostró que, si n es un entero suficientemente grande y si cada uno de los enteros de 1 a n es arbitrariamente impreso de uno de dos colores, existirá siempre una progresión aritmética monocromática compuesta por cierto número de términos prefijado. En 1963, Alfred W. Hales y Robert I. Jewett descubrieron lo que ha mostrado constituir la esencia del teorema de van der Waerden, mientras investigaban el juego de tres en raya. Aunque el tatetí o tres en raya clásico puede acabar aburriendo, el tatetí tridimensional de cuatro en raya resulta endiabladamente difícil. El tablero del juego tridimensional tiene 64 casillas en disposición cúbica. Los jugadores van alternativamente llenando las casillas con ceros y cruces hasta que uno de los jugadores gana por haber ocupado cuatro en línea. El tatetí bidimensional y tridimensional suele terminar en empate. Pero, ¿qué sucederá en juegos de mayor número de dimensiones? ¿Habrá un jugador que pueda asegurarse la victoria en alguna versión n-dimensional de "k-en raya" del juego?

Hales y Jewet demostraron que, si la dimensión n es suficientemente grande, siempre se puede hallar una versión de k-en raya que nunca acaba en empate. Por ejemplo, con independencia de los ceros y las cruces en una versión tridimensional de tres en raya, quedarán alineados tres ceros o tres cruces.

Se puede deducir el teorema de van de Waerden a partir del resultado de Hales y Jewett empleando una transformación que convierte las líneas del tatetí en progresiones aritméticas. Fijémonos en un juego de tatetí en tres dimensiones.



Las coordenadas correspondientes a las cruces de esta combinación ganadora son, 121, 222 y 323, que forman una progresión aritmética. Puede demostrarse que cualquier combinación ganadora transformada por este procedimiento producirá una progresión aritmética.

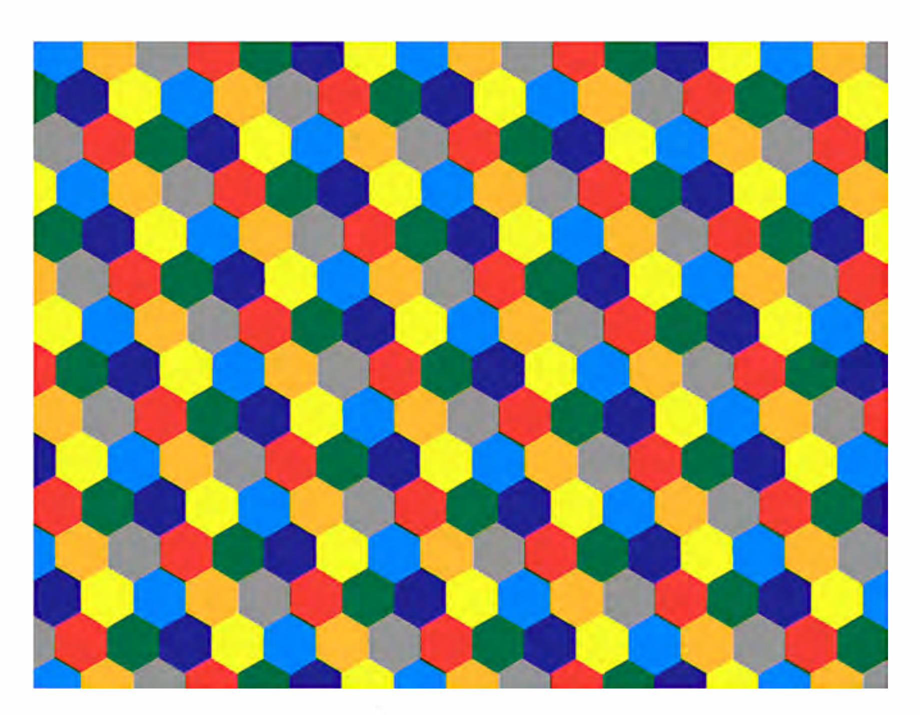
Resultó, sin embargo, que van der Waerden no pudo demostrar el resultado para dos colores sin hacerlo simultáneamente para un número arbitrario de colores.

En su demostración, van der Waerden se valió de una forma especial de inducción matemática. La forma habitual, conocida por inducción simple, consta de dos pasos. Se demuestra primero que el resultado es verdadero para algún valor pequeño, por ejemplo, 2. A continuación se demuestra que, si el resultado es verdadero para un valor cualquiera, también lo es para el valor inmediatamente consecutivo. Se sigue entonces que es verdadero para 3 y, por tanto, para 4, etcétera. El resultado cae como una ristra infinita de piezas de dominó.

Para demostrar el teorema de Ramsey relativo a las progresiones aritméticas, van der Waerden utilizó una inducción doble, más sutil. Supuso que, para cualquier número fijo de colores, existía un número n tal que, si cada uno de los enteros de 1 a n fuese impreso en uno de tales colores, habría necesariamente una progresión aritmética monocromática de, pongamos por caso, 10 términos. Pudo a continuación deducir que, para cualquier número fijo de colores, existía un número m tal que, si cada uno de los enteros de 1 a m fuese impreso en uno de estos colores, habría necesariamente una sucesión monocromática de 11 términos. En general, van der Waerden demostró que saber que el teorema es cierto para k términos y todos los números de colores implica el resultado para $k + 1$ términos y todos los números de colores.

Una vez que van der Waerden hubo llegado a esta fase de la prueba, sólo le faltaba demostrar que el resultado era cierto para algún pequeño valor de k. Si el número de enteros es una unidad mayor que el de colores, habrá siempre dos enteros del mismo color. Estos dos enteros forman una progresión aritmética de dos términos. Así, si el número de enteros supera en una unidad el de los colores, habrá siempre una progresión aritmética cromática de dos términos. El conjunto infinito de dominós con dos términos derriba ahora al conjunto infinito con tres términos, el cual tumba a su vez al conjunto infinito con cuatro términos, y así sucesivamente [*véase el recuadro de la página anterior*].

Habiendo demostrado el teorema de Ramsey para progresiones aritméticas, van der Waerden aplicó ahora su conocimiento al siguiente problema: ¿Cuál es el mínimo valor de n que garantizará una progresión aritmética monocromática de 10 términos (sea por caso), sabiendo que cada uno de los enteros de 1 a n ha sido arbitrariamente impreso de uno de dos colores? La mejor solución que van der Waerden pudo hallar era tan grande que se hace imposible escribirla en la notación habitual. Era mayor que mil millones, mayor que 10 elevado a la potencia mil millones.



4. CIERTAS NOCIONES de la teoría de Ramsey son aplicables a problemas geométricos, como este rompecabezas hexagonal. Si los lados de los hexágonos tienen todos longitud de 0,45 (la unidad es arbitraria), la máxima distancia entre dos puntos un hexágono será 0,9 unidades. Cada hexágono está pintado de uno de siete colores, de manera que no haya dos hexágonos del mismo color que disten menos de 1,19 unidades. Dos puntos del mismo color no están nunca a una unidad. No se ha podido determinar si es posible colorear el plano con seis tintas de forma que ningún par de puntos del mismo color disten exactamente una unidad. (Ilustración preparada por Johnny Johnson.)

De hecho, para expresar este número, los matemáticos recurren a una secuencia de funciones conocida por jerarquía de Ackermann. La primera función de esta jerarquía se llama sencillamente DOBLE (x). Como su nombre da a entender, la función duplica el número x. Así pues, DOBLE (1) es igual a 2, y DOBLE (50) = 100. La segunda función, EXPONENTE (x), puede ser expresada como 2 elevado a la potencia x; consiguientemente, EXPONENTE (3) es igual a 8. También podemos describir EXPONENTE en función de DOBLE. Para hallar EXPONENTE (3), por ejemplo, duplicamos 1, duplicamos el resultado y luego volvemos a duplicar. Cada una de las funciones de la jerarquía de Ackermann está definida en función de su predecesora.

Por tanto, la tercera función de la jerarquía, TORRE (x), es expresable a partir de EXPONENTE. TORRE (3), por ejemplo, es 2 elevado a la potencia de 2 elevado a 2, que es igual a 2 elevado a 4, o sea, 16. TORRE (x) se escribe en ocasiones como una torre de potencias

$$2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}$$

donde el número de doses de la torre es x. Empero, ni siquiera la función TORRE (x) crece con rapidez suficiente para describir el resultado de van der Waerden.

La función siguiente, a la que informalmente se describe como wow (x), se halla comenzando en 1 y aplicando x veces la función TORRE. Así, pues,

wow(1) = TORRE (1) = 2
wow(2) = TORRE (2) = 4
wow(3) = TORRE (4) = 65.536

Para hallar wow (4) se requiere calcular TORRE (65.536). A tal fin hay que comenzar en 1 y aplicar EXPONENTE 65.536 veces. Al aplicar EXPONENTE tan sólo cinco veces se obtiene ya $2^{65.536}$, número cuyas cifras llenarían dos páginas de esta revista. La verdad es que, aunque las cifras de un número llenasen todas las páginas de todos los libros y todos los bancos de memoria de todos los ordenadores, el

número en cuestión sería incomparablemente menor que WOW (4).

Para dar el resultado de van de Waerden es preciso definir una función que crezca más deprisa. La función ACKERMANN (x) se define por la secuencia DOBLE (1), EXPONENTE (2), TORRE (3), WOW (4), y así sucesivamente. ACKERMANN (x) acaba por dominar a la totalidad de las funciones de la jerarquía. La demostración de van der Waerden proporcionó el siguiente resultado cuantitativo: si los enteros 1, 2,..., ACKERMANN (k) se colorean de dos colores, existe siempre una progresión aritmética monocromática de k términos.

Parecía ridículo que un inocente enunciado relativo tan sólo a progresiones aritméticas pudiera originar números tan fabulosamente grandes. A lo largo de los años, los matemáticos se esforzaron por mejorar la demostración de van der Waerden. Al irse sumando los fracasos, comenzó a cobrar cuerpo la idea de que la doble inducción y la correspondiente función ACKERMANN eran características necesarias de cualquier demostración del teorema de van der Waerden. Los lógicos se esforzaron con creciente intensidad en dar argumentos para que realmente fuera así.

En 1987, sin embargo, el lógico israelí Saharon Shelah, de la Universidad Hebrea de Jerusalén, logró un avance importante. Shelah posee dilatada reputación de ser uno de los más vigorosos solventadores de problemas de la matemática moderna. Logró abrirse paso a través de la barrera ACKERMANN y demostró lo siguiente: si los enteros 1, 2,..., WOW (k) se colorean de dos tintas, tiene que haber siempre una progresión aritmética monocromática de k términos.

A pesar de su formación, la demostración de Shelah no utiliza para nada las herramientas de la lógica matemática. Su demostración se vale de ideas matemáticas elementales (aunque superlativamente ingeniosas). Desarrollada por completo, su prueba ocupa unas cuatro páginas, y la mayoría de los especialistas la consideran más clara que la original de van der Waerden. Y lo que es más importante, elude la doble inducción. Fija el número de colores en dos (o en otro número particular) y demuestra seguidamente una inducción simple: si el resultado es verdadero para progresiones de k términos, también lo será para progresiones de ($k + 1$) términos.

Los matemáticos están escrutando la demostración de Shelah y su posible mejora para dar una función TORRE o incluso EXPONENTE para la demostración del teorema de van der Waerden. Uno de nosotros (Graham) ha ofrecido una recompensa de 1000 dólares por una demostración (o refutación) de que, para cada número k, si los números 1, 2,..., TORRE (k) son bicoloreados, ha de formarse necesariamente una progresión aritmética monocromática de k términos.

Los trabajos de Ramsey, Erdös, van der Waerden y de otros muchos establecieron los fundamentos de la teoría de Ramsey. No obstante, los investigadores apenas si han comenzado a explorar las consecuencias de la teoría. Esta sugiere que gran parte de la estructura esencial de las matemáticas consiste en números y conjuntos ingentes, objetos tan grandes que resultan difíciles de expresar y, no digamos ya, de comprender.

Conforme vayamos aprendiendo a manejar estos números grandes, tal vez podamos hallar relaciones matemáticas que ayuden a los ingenieros a construir complejas redes de comunicaciones o que faciliten a los científicos el reconocimiento de regularidades en sistemas físicos de gran escala. En nuestros días podemos fácilmente darnos cuenta de que las constelaciones del firmamento nocturno son consecuencia de la teoría de Ramsey. ¿Qué regularidades podremos hallar en conjuntos que sean ACKERMANN (9) veces mayores?

BIBLIOGRAFIA COMPLEMENTARIA

COMBINATORIAL RELATIONS AND CHROMATIC GRAPHS. A. M. Gleason y R. E. Greenwood en *Canadian Journal of Mathematics*, vol. 7, n.º 1, págs. 1-7; 1955.

HOW THE PROOF OF BAUDET'S CONJECTURE WAS FOUND. B. L. van der Waerden en *Studies in Pure Mathematics*. Dirigido por L. Mirsky. Academic Press, Inc., 1971.

PAUL ERDÖS: THE ART OF COUNTING: SELECTED WRITINGS. Dirigido por Joel Spencer. The MIT Press, 1973.

THE MAN WHO LOVES ONLY NUMBERS. Paul Hoffman en *Atlantic Monthly*, vol. 260, n.º 5, págs. 60-74; noviembre de 1987.

NUMBERS IN RAMSEY THEORY, IN SURVEYS AND IN COMBINATORICS. R. L. Graham y V. Rödl. Conferencias de la Sociedad de Matemáticas de Londres, n.º 123, págs. 111-153; 1987.

RAMSEY THEORY. Ronald L. Graham, Bruce L. Rothschild y Joel H. Spencer. Segunda edición. John Wiley & Sons, Inc., 1990.

Juegos de ordenador

Un sin par viaje por calles pares nos trae de vuelta a casa en Golípolis

A. K. Dewdney

"Un viaje de mil millas ha de empezar por un paso."

—Lao-Tsé,
La senda de Lao-Tsé

Permita que le inicie en los recorridos de Golípolis. Puede usted dar paseos parecidos en Barcelona, Nueva York o cualquier gran ciudad cuyas calles formen un cuadriculado. He aquí las instrucciones. Camine calle abajo una manzana; al llegar al cruce, tome a la derecha o a la izquierda, avance en línea recta dos manzanas esta vez, siga a la derecha o a la izquierda, avance tres manzanas, y así sucesivamente. Siempre que doble una esquina, tiene que caminar en línea recta una manzana más que la vez anterior. Si al cabo de cierto número de giros retorna usted al punto de partida, habrá descrito un golígono. Y en caso de que no desee usted hacer ejercicio, el paseo se puede simular sin dificultad moviendo un lápiz sobre papel cuadriculado. Si se pierde usted, sírvase consultar el plano de la figura 1.

Un golígono está formado por segmentos rectilíneos cuyas longitudes (medidas en metros, kilómetros o lo que se quiera) creen la secuencia 1, 2, 3, etcétera, hasta un número finito. Cada segmento conecta en ángulo recto con el segmento que es una unidad mayor, a excepción del más largo, que engarza con el primero (y más corto) en ángulo recto. Los golígonos no constituyen mera curiosidad geométrica. Han inspirado rompecabezas deliciosos, amén de problemas de investigación desconcertantes. ¿Y qué mejor manera de hacernos idea de los procesos de investigación que llevar a cabo un viaje recreativo?

Parece que los golígonos fueron concebidos por Lee Sallows, un formidable ingeniero de la Universidad Católica de Nimega, en Holanda. La última vez que expuse un trabajo de Sallows fue en diciembre de 1984, fecha en que describí su búsqueda de pangramas, frases que contienen todas las letras del alfabeto. Desde aquella fecha, Sallows ha inventado muchas recreaciones nuevas, pero ninguna tan cautivadora, en lo tocante a nuestra sección, como los golígonos.

Sallows comenzó a buscar golígonos en otoño de 1988. No tardó mucho en encontrar un golígono de ocho lados. No logró, empero, hallar objetos de esta clase que tuvieran menos de ocho lados, ni pudo tampoco descubrir ninguno con 9, 10 u 11... hasta que se topó con uno de dieciséis.

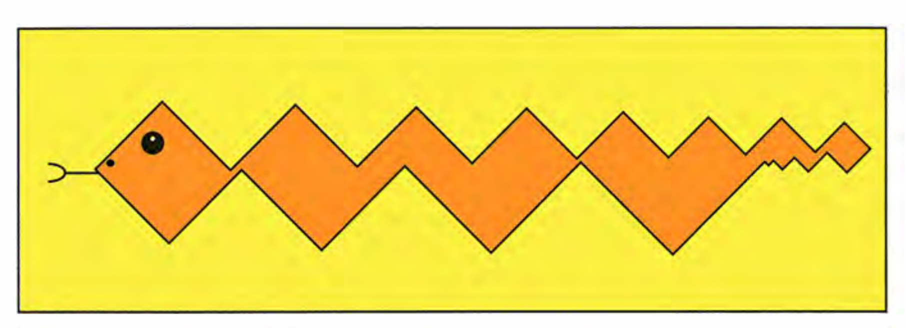
Preguntándose si se habría saltado alguno, Sallows redactó un programa de ordenador para automatizar la búsqueda. El programa engendró no menos de 28 golígonos diferentes de 16 lados [*véase la figura 2*] antes de proseguir hacia órdenes más elevados. El programa no encontró golígonos de 17 a 23 lados, pero sí generó muchísimos de 24, concretamente, 2108, por ser exactos.

Para entonces, Sallows tenía la corazonada de que el número de lados de un golígono había de ser múltiplo de ocho. Sin embargo, su programa, que estaba trabajando ya en los veintitantos lados, no lograba hallar golígonos de 32. Desilusionado, recurrió a Martin Gardner, autoridad suprema en cuestiones de matemáticas recreativas. ¿Podría Gardner demostrar que el número de lados de un golígono debía ser múltiplo de 8?

Podemos ir escalando la demostración de Gardner partiendo de la tarea, más sencilla, de probar que el número de lados de un golígono tiene que ser múltiplo de dos. Para explicarla, permítame usted llevarle de vuelta a Golípolis. Iniciamos nuestro recorrido caminando una manzana en dirección norte. (He elegido el norte por pura conveniencia.) Por tanto, el primer lado del golígono subirá una manzana hacia el norte. A continuación, podemos doblar hacia la derecha o la izquierda, esto es, hacia el este o el oeste. En consecuencia, el segundo lado del golígono avanzará dos manzanas hacia el este o dos manzanas hacia el oeste. Así, pues, todos los lados impares del golígono (el primero, el tercero, el quinto...) son de longitud impar y corren hacia el norte o hacia el sur; todos los lados pares (el segundo, cuarto, sexto, etcétera) poseen longitudes pares y corren ha-



1. Un plano de Golípolis.



2. *Una sierpe goligonal de 32 lados y lista completa de los 28 golígonos de 16 lados.*

cia el este o el oeste. Puesto que el último lado forma ángulo recto con el primero, el último lado ha de ir hacia el este o el oeste. El último tramo corresponde a un lado par del golígono y el número de lados del golígono tiene que ser múltiplo de 2.

Para demostrar que el número de lados ha de ser múltiplo de cuatro empezamos por calcular la distancia total que hemos viajado hacia el norte del punto de partida. Basta para ello sumar el número de manzanas recorrida en dirección norte y restarle el número de manzanas recorridas en dirección sur. (Si la distancia neta final resulta negativa, ello significa, sencillamente, que nos encontramos al sur del punto de partida.)

Dado que todos los tramos en dirección norte o sur tienen longitud impar, lo que estamos haciendo es, en esencia, sumar y restar números impares consecutivos. Unos sencillos tanteos y razonamientos nos convencerán de que, al sumar o restar un número par de recorridos nones, el resultado siempre es par; por ejemplo: $1 + 3 - 5 + 7 = 6$. Por la misma razón, un número impar de longitudes nones dará siempre suma impar. Así, pues, la distancia neta recorrida en dirección norte será de un número par de bloques, si y solamente si hemos caminado a lo largo de un número par de lados norte o sur.

Ahora, si caminamos por Golípolis desde un punto y retornamos a él, la distancia hacia el norte del punto de partida es igual a cero. Dado que cero es número par, el golígono ha de tener un número par de lados que corran hacia el norte o el sur. El número total de lados es el doble del número de lados norte o sur, ya que por cada lado norte o sur ha de haber un lado este u oeste. Por tanto, el número de lados de un golígono es un múltiplo de cuatro.

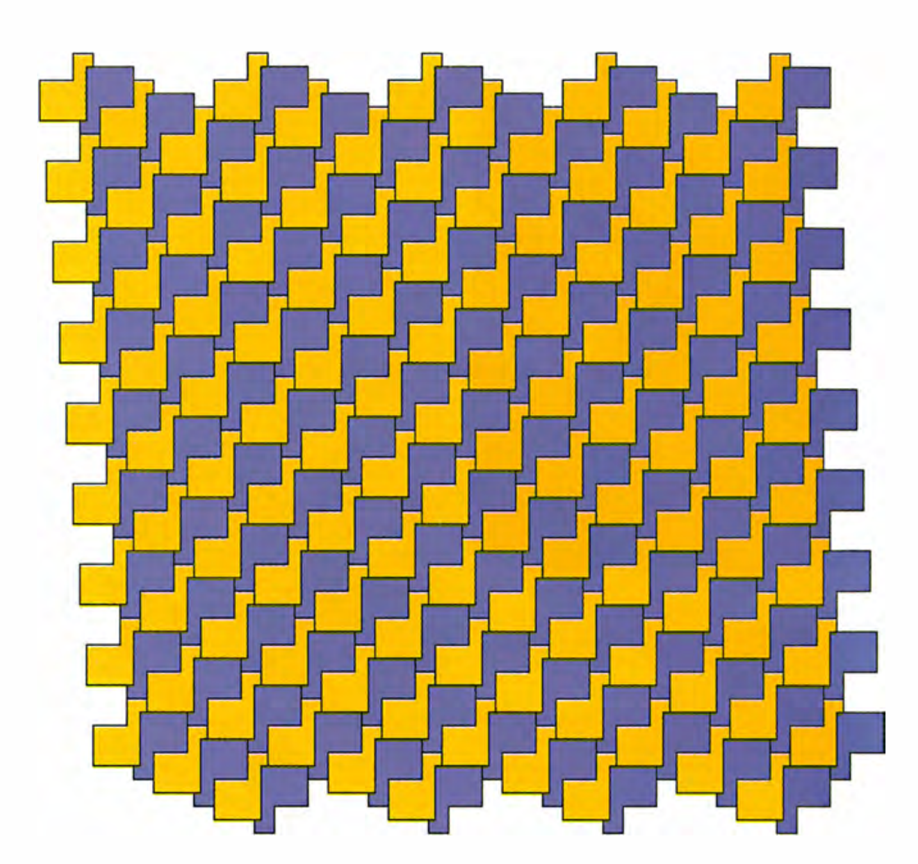
¿Y cómo se las arregló Gardner para demostrar que el número de lados ha de ser un múltiplo de ocho? Acompañemos al maestro en su periplo. Sabemos ya que el número de lados de dirección general norte-sur es par. Sabemos también que la distancia avanzada hacia el norte tiene que ser igual a la distancia recorrida hacia el sur. Por consiguiente, el recorrido hacia el norte tiene que ser la mitad de la distancia total caminada en ambas direcciones, norte y sur. Y de aquí, la distancia total es igual a la suma de un número par de números impares consecutivos. Por ejemplo,

$$1+3=4$$
$$1+3+5+7=16$$
$$1+3+5+7+9+11=36$$
$$1+3+5+7+9+11+13+15=64$$

El resultado parece ser siempre un múltiplo de cuatro.

Si la distancia total hacia el norte y el sur es igual a un múltiplo de cuatro, la mitad de esa distancia tiene que ser múltiplo de dos. Por consiguiente, la distancia que hemos viajado hacia el norte tiene que ser múltiplo de dos, y el número de lados que corren hacia el norte tiene también que ser múltiplo de dos, porque todos ellos tienen longitudes impares. Se sigue que el número de lados hacia el sur es múltiplo de dos y que el total de lados que van sea al norte, sea al sur, es múltiplo de cuatro. Finalmente, después de contar también los lados hacia el este y el oeste, el número total de lados del golígono tiene que ser un múltiplo de ocho.

La demostración de Gardner puede parecernos bastante peculiar, porque en ella los números impares desempeñan papel fundamental, mientras que los pares apenas si intervienen en el análisis. Sin embargo, se puede llegar a la misma conclusión efectuando un razonamiento similar basado en longitudes pares. En realidad, una y otra clase de longitudes desempeñan papeles independientes en la construcción de cualquier golígono. Así, por ejemplo, podemos obtener un nuevo golígono a partir de dos cualesquiera de igual tamaño utilizando las longitudes pares del primer golígono y las impares correspondientes al segundo.



3. El golígono mínimo tesela el plano.

Hasta ahora, nuestra senda recreativa nos ha llevado desde el problema de existencia a la cuestión de las condiciones de existencia; definitivamente, el número de lados de un golígono ha de ser necesariamente múltiplo de ocho. Por comodidad, llamaré a este teorema condición $8k$ para golígonos bidimensionales. El teorema enuncia lo que los matemáticos llaman "una condición necesaria": si existe un golígono, el número de sus lados es necesariamente múltiplo de ocho. Pero, ¿será suficiente esta condición? Con otras palabras, si meramente pedimos que el número de lados, n, sea un múltiplo de ocho, ¿será ello suficiente para garantizar la existencia de un golígono?

Podemos hallar un golígono cuyo número de lados sea un múltiplo arbitrario de 8 obedeciendo a un sencillo juego de instrucciones. El primer paso es determinar cuán grande queremos que sea n. Para ilustrar el método me decidiré por un golígono de 16 lados. El segundo paso consiste en escribir los números consecutivos de 1 a n, en este caso, de 1 a 16.

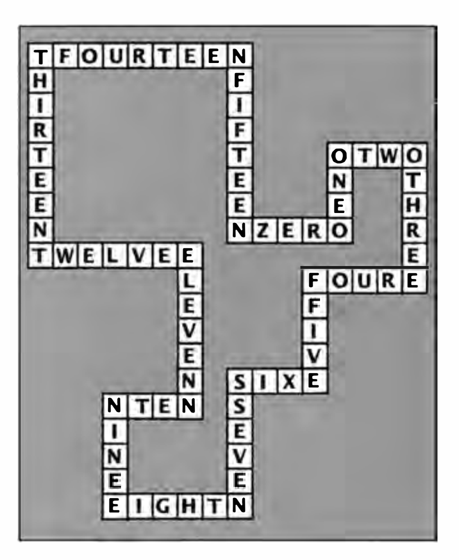
El tercer paso ha de ser dotar a estos números de signos positivos y negativos. La primera y la última cuarta parte de la lista recibirán signos positivos, y todos los números intermedios, signos negativos: +1, +2, +3, +4, −5, −6, −7, −8, −9, −10, −11, −12, +13, +14, +15, +16. Para traducir esta secuencia de números en un periplo por Golípolis basta saber que los números impares positivos corren hacia el norte, los impares negativos, al sur, los pares positivos hacia el este, y los pares negativos, al oeste. El golígono correspondiente se asemeja a una sierpe. Vemos una "sierpe goligonal" de 32 lados en la figura 2.

Tras dejar zanjada la cuestión de la suficiencia de la condición $8k$, Gardner y Sallows comenzaron a preguntarse cuántos golígonos existirían. Habían hallado uno de orden 8, 28 de orden 16, y 2108 de orden 24. ¿Cuántos podrían existir de órdenes más altos? Quienes más suelen formularse preguntas de este tipo son los matemáticos especializados en combinatoria, ciencia que estudia objetos matemáticos discretos como los conjuntos finitos y los grafos [véase "Teoría de Ramsey", por Ronald L. Graham y Joel H. Spencer, en este mismo número]. Gardner decidió consultar a algunos expertos. Envió cartas al científico informático Donald E. Knuth, de la Universidad de Stanford, y al matemático Richard K. Guy, de la Universidad de Calgary, en Alberta. ¿Podría alguno de ellos ayudarle en el problema de la enumeración de los golígonos?

Antes de mucho, Knuth había preparado un programa de ordenador que contaba todos los golígonos de hasta 64 lados. El número total de golígonos de 64 lados, por ejemplo, es 127.674.038.970.623. Desdichadamente, el programa de Knuth dejaba de ser práctico para valores superiores.

Guy, entretanto, logró idear una fórmula que expresaba el número aproximado de golígonos para cada posible valor de k, siendo k igual al número de lados dividido por ocho. La fórmula de Guy consiste en una fracción, cuyas partes principales son, en el numerador, 2 elevado a $8k$-6, y k elevado al cubo en el denominador. Dado que al crecer k el término exponencial aumenta mucho más rápidamente que el término cúbico, el crecimiento del número de lados es fundamentalmente exponencial. La fórmula de Guy posee además la importante propiedad adicional de ser asintótica, es decir, da un valor que expresado como tanto por ciento del valor verdadero tiende más y más al 100 por ciento conforme k crece y crece.

Sallows, Gardner, Knuth y Guy decidieron dar cuenta en un artículo de sus diversos hallazgos. La palabra "golígono" no figura en el título. El artículo se llama, en cambio, "Serial



4. Un logolígono creado por Sallows.

Isogons of 90 Degrees", nombre de más solemnes resonancias que "golígonos". La alusión a los 90 grados lleva implícito que se podrían ensamblar golígonos de otras clases. De hecho, algunos de los autores han examinado golígonos cuyos lados no forman ángulos de 90 grados, sino de 60 o 120. Para construir tales golígonos, los lectores pueden servirse de papel hexagonal o isométrico, que se puede hallar en tiendas de suministros de delineación.

¿No habría sido éste lugar adecuado para poner el punto final? La investigación matemática (lo mismo que la matemática recreativa) no se detiene nunca. Con frecuencia, al progresar la investigación en una cuestión se nos plantean otras cuestiones. A veces resulta muy difícil decidir exactamente en qué cuestión vamos a trabajar.

Apenas han sido resueltos los problemas de investigación, y ya está el matemático recreativo que llevamos dentro pensando en la indagación de nuevas avenidas. ¿Podríamos hallar golígonos cuyos lados fuesen números primos? En este caso, las longitudes de los lados consecutivos habrían de ir aumentando según la serie de los números primos impares: 1, 3, 5, 7, 11, 13, y sucesivos.

Es digno de mención otro de los retoños de los golígonos. Se insta a los lectores a examinar nuevamente el golígono de ocho lados. No hay otro, es el único miembro de su clase. Los "octolígonos" se acoplan perfectamente unos con otros y permiten crear motivos para pavimentos. Como dicen los profesionales, "teselan el plano". La figura 3 muestra una atractiva pavimentación.

Hablando con propiedad, la región limitada por un golígono de ocho lados debería llamarse poliominó. No todos los golígonos producen poliominós, porque algunos de ellos se cortan a sí mismos en su sinuosa marcha por el plano. De hecho, es probable que los golígonos poliomínidos (formidable expresión gracias a la cual pueden los lectores perder amigos y tratar de impresionar al personal) se vayan haciendo cada vez más escasos en términos relativos al crecer *n*. Solamente cuatro de los 29 golígonos mostrados en la figura 2 son contornos de poliominós. Al parecer, ninguno de ellos puede teselar el plano.

Pero en materia de teselaciones, Sallows tiene un problema con que retar a los lectores: supongamos que el octolígono represente el contorno del suelo de nuestra cocina. Dados 13 perfiles en forma de L, ¿podrán los lectores recubrir exactamente el suelo de la cocina con estas disparatadas losetas? Para esta cuestión, podemos imaginar que el suelo está dividido en 52 cuadros, con los cuadrados del papel especial en que dio comienzo toda nuestra indagación. El lado de cada cuadrado mide un paso. Los perfiles en L constan de tres cuadrados en línea más otro en ángulo recto, que forma el pie de la L. Obviamente, quienes consideren que la tarea es imposible han de probar que así es.

Concluiré con unas palabras de Sallows, el ingeniero que dio comienzo a todo esto. Las palabras son ZERO, ONE, TWO, y así sucesivamente, hasta FIFTEEN (cero, uno, dos,..., quince). Resulta posible enlazarlas a todas en un golígono especial, en el cual las letras de estas palabras determinan los lados de lo que únicamente podríamos llamar "logolígono".

La caja de Pandora de mentes, máquinas y metafísica que destapé el pasado mes de marzo me ha facilitado silla de pista desde la que presenciar un combate entre titanes. En un rincón, Roger Penrose, cuyo reciente libro, *The Emperor's New Mind*, ha infundido renovados alientos a quienes consideran que el cerebro humano es mucho más que una computadora. En el rincón opuesto, miembros de la comunidad informática y de inteligencia artificial, que no ven en el libro de Penrose nada particularmente nuevo e incisivo.

En ayuda de Penrose fue requerido el filósofo John R. Searle [véase "¿Es la mente un programa informático?", INVESTIGACIÓN Y CIENCIA, marzo de 1990]. Searle expuso el argumento de la "sala china": un programa de ordenador que sea capaz de leer y escribir en chino no tiene del chino mayor comprensión que una persona que imite la acción del programa. Entre los paladines de la inteligencia artificial se cuenta Hans P. Moravec, experto en robótica de la Universidad Carnegie-Mellon, quien me envió una carta abierta dirigida por él a Penrose. Moravec, cuya actitud apasionada y agresiva queda abundosamente manifiesta en su libro *Mind Children*, dice que la argumentación de Penrose es de "obstinación cerril".

¿Estaría Penrose dispuesto a conceder a un cierto robot la misma inteligencia que le atribuye al perro que le mira agradecido con sus ojazos pardos o le gime quejumbroso? El robot en que Moravec está pensando es hasta el presente más imaginario que real. El robot dice, "Penrose, por favor, me duele mucho que no me consideres una persona. ¿Qué puedo hacer para convencerte? Yo tengo conciencia de ti y tengo conciencia de mí. Y he de decírtelo francamente, tu actitud me resulta casi insoportable". La reacción del robot es resultado de un programa especial concebido para convencer al amo de que merece un poco de atención —algo así, quizá, como los gemidos del perro.

Apenas puedo empezar a hacer un muestreo de los muchos lectores que respondieron a aquel artículo. Jef Raskin, de Pacifica, California, afirma que el argumento de la sala china de Searle "equivale a definir la inteligencia declarando que la inteligencia es un proceso no algorítmico que...". Paul Preuss, de San Francisco, aficionado a la informática y distinguido autor de fantasía científica, afirma que la sala china no constituye un auténtico experimento conceptual. Tendría que ser realizado, dice Preuss, "por el Dr. Searle en persona. En el caso de que lograse convencer a un chino parlante de que sabe leer y escribir el chino, debería dar cuenta de sus juicios subjetivos acerca de si durante el proceso ha aprendido chino o no".

Es corriente citar el famoso teorema de incompletitud del matemático alemán Kurt Gödel como prueba de que existen pensamientos que trascienden de lo alcanzable por vía algorítmica. Mi pregunta de si el teorema goza de rango extra-algorítmico produjo un clamor de "noes".

BIBLIOGRAFIA COMPLEMENTARIA

SERIAL ISOGONS OF 90 DEGREES. Lee Sallows, Martin Gardner, Richard Guy y Donald Knuth en *Mathematics Magazine*. Dirigido por G. L. Alexanderson (en prensa).

Juegos matemáticos

Representación matematizada de las especies, de sus aptitudes y del curso de su evolución

Ian Stewart

Una de las controversias más exaltadas de la actualidad —y digo una, pues hay otras varias, entremezcladas— en el campo de la evolución es la que enfrenta a los que podríamos denominar puntualistas y gradualistas. Los gradualistas sostienen que la evolución es proceso continuo que procede a pasitos diminutos. Los puntualistas afirman que la evolución de especies nuevas sólo se produce en rápidos saltos. Todos hemos visto en los textos escolares esos elegantes árboles de la evolución, cuyas ramas se bifurcan y se alzan majestuosas hacia el cielo, para culminar triunfalmente en el *Homo sapiens*, encaramado en lo más alto como el cometa de un árbol de Navidad. Estos árboles reconstruidos quedan muy pulcros, pero la realidad es que los registros fósiles en que se basan son mucho más irregulares; con frecuencia se asemejan mucho más a la figura 1. Los gradualistas consideran que estas aparentes discontinuidades no son más que lagunas en la serie fósil, que serán quizás eliminadas en el futuro por el descubrimiento de fósiles nuevos, o debidas a la errática manera en que sedimentan las rocas. Es preciso decir que su explicación es razonable. Los puntualistas, por su parte, opinan que las discontinuidades del registro fósil son genuinas, y que constituyen, de hecho, la única fuente de especies nuevas. ¿Quiénes están en lo cierto, los puntualistas, los gradualistas... o ninguno de ellos?

No hace mucho, descansaba yo en el sofá, con el gato en el regazo, meditando perezosamente sobre el problema de los cambios graduales y los cambios repentinos en la evolución... cuando llamaron a la puerta. Era el cartero, que me traía un paquete. El paquete, envuelto en papel marrón liso, llevaba adherido un marbete azul con la inscripción "EcoMat de Instrumentación, S. L.", y otro rojo donde se leía "¡MUY FRAGIL!". Me quedé un poco cortado, porque jamás había oído hablar de tal compañía y mucho menos les había encargado nada. Mas la especulación basada en el aspecto externo de paquetes sellados es empresa vana, así que retiré la envoltura y hallé una tarjeta rotulada "Evoluscopio Zeeman. Prototipo 1". [Así llamado en honor de Christopher Zeeman, en cuyo modelo matemático del proceso evolutivo se inspira este artículo.]

Picada mi curiosidad en lo más vivo, abrí la caja y descubrí un extraño artilugio a modo de casco, una especie de híbrido entre minicadena estereofónica y de gafas de sol. La caja de control se hallaba cubierta de diales, botones, mandos, palancas y diodos luminiscentes. Sí, eso era justo lo que parecía, una minicadena.

Un grueso manual de anillas y cubiertas plastificadas yacía bajo ella.

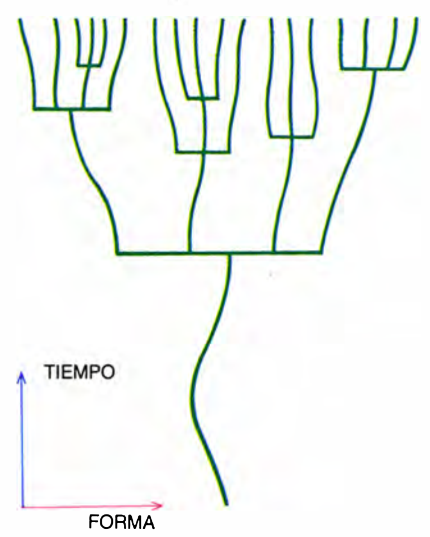
Extraje el aparato y me lo coloqué en la cabeza. Tuve una extraña sensación de discontinuidad; la sala se hizo aguas por un instante; después volvió a enfocarse y todo se tornó diferente. En lugar de habitación había un amplio espacio llano. En realidad, no era verdaderamente un plano: cuanto más me fijaba, más dimensiones alcanzaba a ver: cientos, miles, millones... Pero yo seguía considerándolo un plano. Estaba poblado por las criaturas más curiosas que jamás hubiera visto, densamente apiñadas [*véase la figura 2*]. Bajé la mirada hacia el lugar donde habría de hallarse mi regazo, y vi en él a mi gato, durmiente ovillo de pelaje marrón. Pero a su lado había otro gato, y otro, y otro —millones de gatos, más empaquetados que sardinas... exhibiendo sutiles diferencias–. Los había atigrados, los había blancos y negros, gordos y lustrosos unos, flacos los otros. Y al desviar la mirada hacia un lado vi gatos con dos colas, con cinco patas, con tres ojos; vi un gato verde con patines en lugar de garras, un gato transparente, un gato esférico que rodaba por el suelo persiguiendo a un ratón igualmente esférico...

¿Cómo podían caber todos en un espacio tan reducido? Perplejo, eché mano del manual, y con dificultad —pues su redacción era tan buena como la de casi todos los manuales, vale decir, impenetrable— hallé una entrada que hacía al caso.

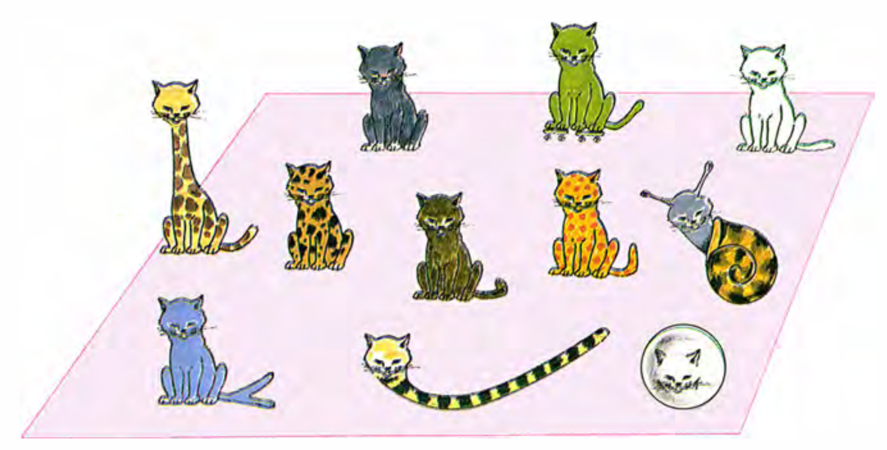
Espacio de formas: El evoluscopio Zeeman ofrece, en su modalidad implícita, una panorámica estática del *espacio de formas* real. Cada punto de este continuo multidimensional corresponde a la forma de una criatura potencialmente viva. El evoluscopio permite la visión directa no sólo del espacio, sino también de las criaturas asociadas a cada uno de sus puntos. Para seleccionar la capacidad de supervivencia, se ha de pulsar en la unidad de control el botón rotulado *función de adecuación*.

Hallé el botón y lo pulsé. El plano empezó a abombarse y a abultarse, dando lugar a un paisaje de redondeados altillos [*véase la figura 3*]. Mi gato flotaba delante de mi nariz, encaramado a una prominencia no muy alta. Había en lontananza perros, una vaca, una jirafa y una lombriz de tierra, subidos cada uno a su altozano. No faltaban otras criaturas extrañas; a una de ellas la bauticé *telefante*, pues parecía un elefante con la trompa en forma de teléfono. Había también una *cabricol* (una especie de cabra provista en el lomo de una concha espiral), una *chimpancebra*, un *pelicanguro* y un *faisantílope* ...

Busqué con la vista el gato a patines, sin éxito al principio, hasta que hundí la mirada en un valle profundo y apartado, donde estaba merodeando, rodeado por otros de su especie. Pronto me di cuenta de que las cria-



1. Discontinuidades en el registro fósil.



2. Una pequeña región del espacio de formas. Hay una criatura en cada punto, cuya forma varía continuamente al moverse el punto.

turas más familiares parecían estar ocupando las cimas de los cerros, y que, en cambio, los animales más insólitos habitaban el fondo de los valles. Al ir mi ojo explorando las laderas desde el pie hacia la cumbre, las criaturas iban gradualmente tornándose más familiares; las cimas estaban pobladas exclusivamente por animales que yo podía reconocer como habitantes normales de nuestro planeta, a excepción de unas pocas colinas, a primera vista desocupadas.

Fui hojeando el manual, y empezó a hacerse la luz.

Función de adecuación. Se obtiene pulsando el correspondiente botón en la unidad de control. Ofrece una medida de la adecuación de cualquier criatura dada del espacio de formas, comprimida esquemáticamente a una magnitud escalar, representada por una *altura.* La adecuación viene representada por una gráfica de esta magnitud en función de los puntos del espacio de formas, lo que tiene por efecto desplazar el espacio de formas en la dirección vertical en proporción a la función de adecuación: en téminos matemáticos, corresponde a mostrar la *gráfica* de la función de adecuación. El plano inicial correspondiente al espacio de formas se ondula y alabea, y ahora representa el *espacio de competencia.* Las especies más aptas para sobrevivir ocupan niveles más altos en este paisaje.

Opción zoom. Haciendo girar el dial situado en el puente de la pieza de la nariz, se puede variar la amplificación a escala local, para hacer visibles los seres microscópicos.

¡Por eso había altillos desocupados! Hice zoom sobre unos cuantos y encontré una ameba, una hidra y una bacteria *Salmonella.*

Ahora sabía ya en qué consistía el evoluscopio. Era una máquina que tornaba visibles los diversos ingredientes matemáticos de la evolución darwinista. Aunque Darwin dio a sus teorías forma verbal, es posible traducirlas a forma matemática, con lo que resulta más fácil analizar sus consecuencias. Evidentemente, aunque cualquier modelo matemático de la evolucion darwinista no pasará de ser una aproximación de la situación real —suponiendo que exista una situación real, lo cual, es desde luego, problema tan controvertido como todo lo que atañe a esta materia—, lo cierto es que una formulación matemática precisa puede ayudarnos a distinguir entre diversas interpretaciones. Estas formulaciones pueden también dejar a la vista incoherencias en el análisis lógico, actuando a modo de "contraejemplos" de ciertos tipos de argumentación verbal. Lo malo de las palabras es que pueden variar sutilmente de significado en el desarrollo de un razonamiento aparentemente lógico, y ello puede invalidar las conclusiones. Las matemáticas son más precisas, y nos obligan a seguir las hipótesis que se han hecho... quizás a expensas de la imaginería poética.

En el modelo particular de evolución de que estaba provisto el evoluscopio, cada forma de vida animal estaba representada por un punto del espacio de formas. Contenía este espacio no sólo seres viables, sino la totalidad de las *formas imaginables.* Disponía igualmente de una medida de la adaptación o aptitud para la supervivencia de cada una de ellas, la función de adaptación; los seres mejor dotados para competir con éxito eran los de nivel de adaptación máximo posible. A los máximos locales de la función de adaptación —las cimas de las colinas del paisaje o espacio competitivo— les correspondía una especie.

Este hecho ayuda a comprender el motivo de que las especies tiendan a estar claramente separadas unas de otras, no obstante la continuidad de la gama de formas posibles. Los ratones, por ejemplo, son francamente pequeños; no se observa una gama continua y extensa de tamaños en las criaturas ratoniles. El espacio de formas es continuo, pero cada máximo local está rodeado por su propio montecillo, y estas elevaciones están separadas unas de otras; tanto los ni-

ADAPTADO
MAXIMO LOCAL DE LA FUNCION DE ADAPTACION
INADAPTADO
OTRO MAXIMO LOCAL
OTRO MAXIMO LOCAL
ESPACIO DE COMPETENCIA
FUNCION DE APTITUD
ESPACIO DE FORMAS

3. El paisaje competitivo. Las formas vivas que ocupan máximos locales son idóneas para competir y sobrevivir; las situadas en los valles no están bien adaptadas y se extinguen.

chos como las especies que los ocupan forman conjuntos discretos.

Otro de los fenómenos que encajan bien con el modelo es la *evolución convergente* —la aparición de formas similares, aunque dotadas de genes completamente distintos. Existe, por ejemplo, un paralelismo sorprendente entre los marsupiales australianos y los mamíferos placentarios de los otros continentes [*véase la figura 4*]. Tenemos otro ejemplo en el cuerpo ictiforme de los delfines.

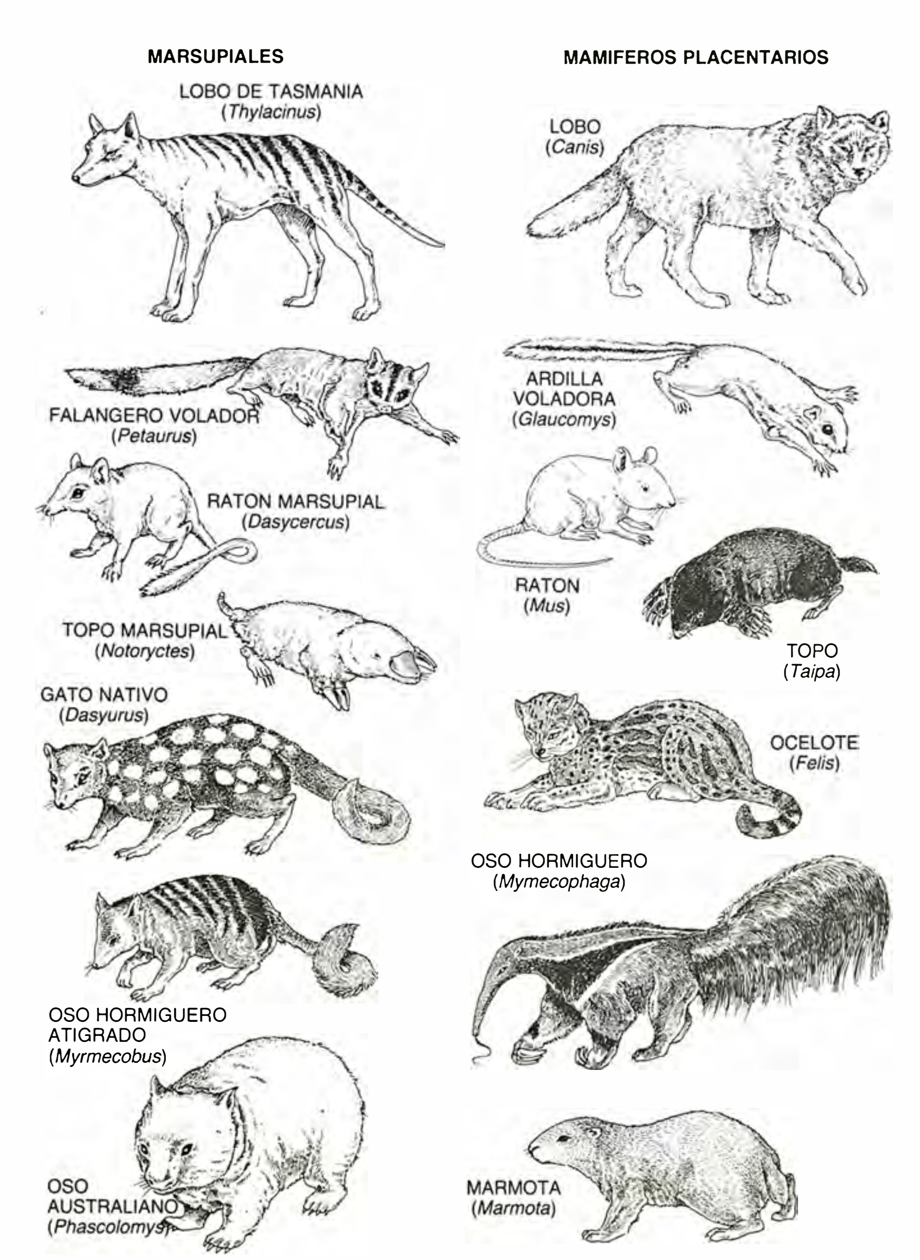
Antes de seguir, estudié el manual de cabo a rabo. Al fin, convencido de que se trataba de una mera simulación, y de que nada de cuanto pudiera hacer iba a afectar el medio ambiente actual de la Tierra, comencé a experimentar. La siguiente entrada del manual iba directamente al meollo del razonamiento de Darwin.

Para disponer de una representación más exacta de la auténtica situación, haga girar el mando rotulado *mutación aleatoria*. ¡AVISO! Si el mando se lleva a la posición "nivel alto", puede producirse caos. Experimente haciendo girar el mando *gradualmente*.

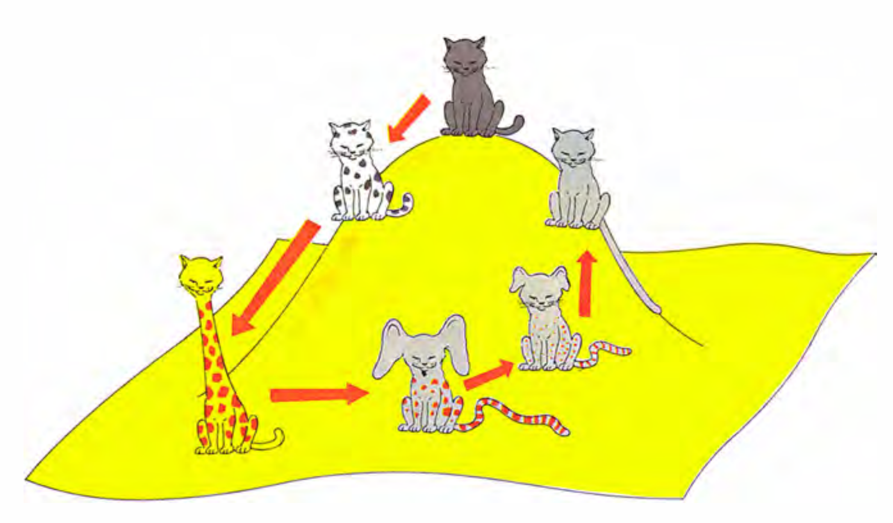
Giré el mando suavemente, y mi gato se tornó pardo, después atigrado, al poco, siamés; por último, la cola le desapareció. Giré el mando un poco más y el gato se tornó borroso. Estaba *vibrando*.

Accioné la palanca de *avance lento* para observar con mayor claridad lo que ocurría. Cada pocos segundos, mi gato cambiaba ligeramente de forma. Y al tiempo que cambiaba, se desplazaba un poco por el espacio de competencia. En el cerrillo se encontraban apiñados toda suerte de gatos, que iban cambiando de forma y lugar en una danza aturullante [*véase la figura 5*]. Un cambio grande ocasional desplazaba al gato más allá de lo normal, pero de ordinario volvía a ascender progresivamente hacia la cima de su cerro. Por ejemplo, durante un cierto tiempo le crecieron unas grandes orejas, como de elefante, pero me fijé en que entonces descendió un poco colina abajo. Era de presumir que la mutación a orejas de elefante le resultase desfavorable, haciendo al gato menos capaz de competir. El puro esfuerzo de ir tirando de las orejas le frenaría considerablemente, y los ratones podrían oír el batir de las orejotas al ir el gato acercándose subrepticiamente a ellos.

Las mutaciones que producían gatos situados cada vez más bajos en la ladera no eran duraderas. Los gatos, aunque sin rumbo fijo, tendían a subir otra vez, retornando a la forma felina más normal, conforme la selec-



4. *Evolución paralela de los marsupiales y de los mamíferos placentarios.* [De la Encyclopaedia Britannica].



5. *Danza aleatoria de los gatos mutantes.*

ción natural iba eliminando las formas menos ortodoxas.

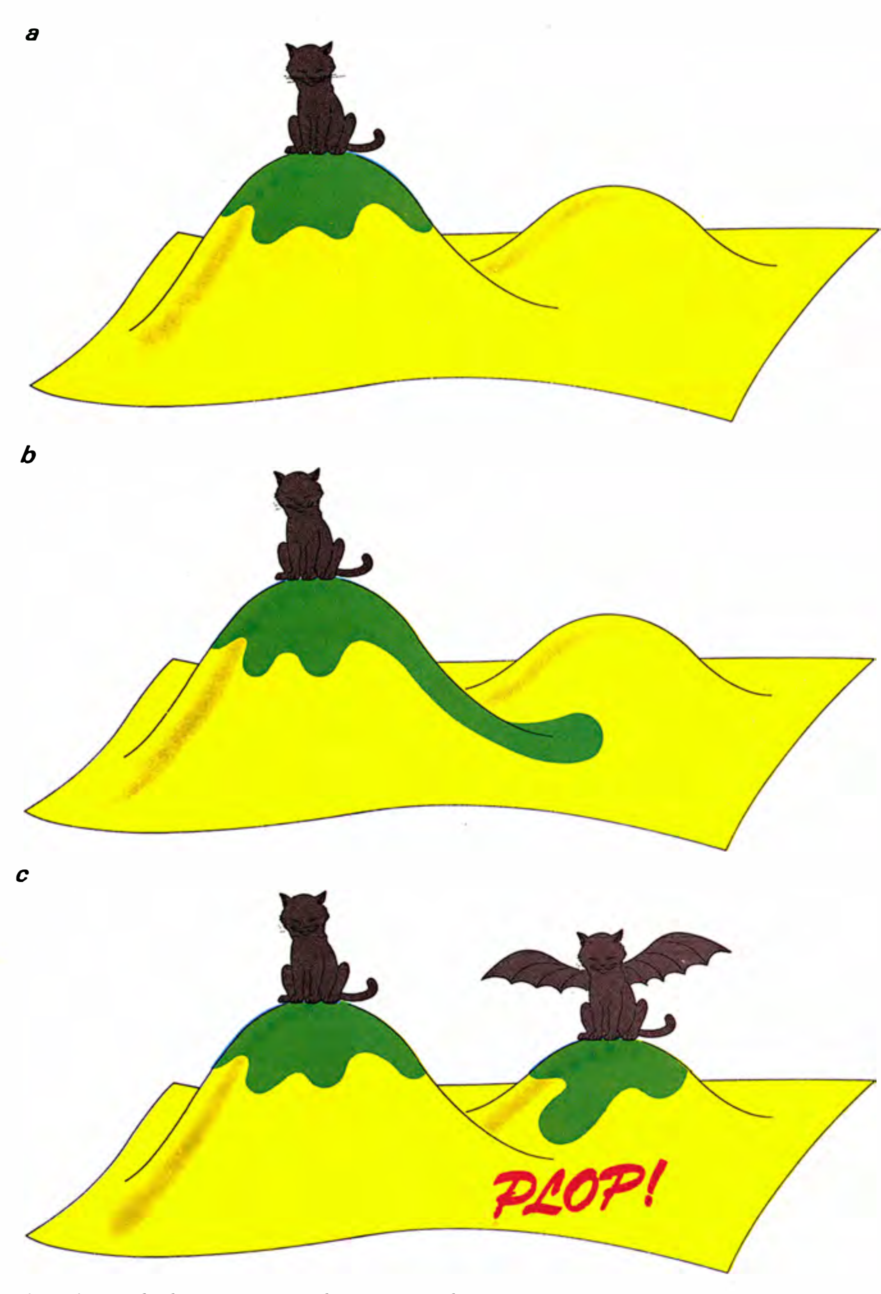
Giré más todavía el mando de *mutación*. El altillo bullía de actividad, surgiendo por todo él gatos en vibración a modo de olas en un mar tormentoso. De repente, con un ¡plop! audible, una especie de glóbulo lleno de gatos se deslizó colina abajo, osciló dubitativo en el valle y luego salió disparado hacia la cumbre de un cerrillo vecino, vacante hasta entonces [*véase la figura 6*].

Volví a poner la "cámara lenta" e inspeccioné a los recién llegados.

Fascinante: El primer gato volador del mundo.

Comprendí entonces lo que había presenciado: *la evolución de una especie nueva*.

Mutaciones aleatorias, de frecuencia muy elevada, comprimidas en unos pocos minutos, habían producido una forma nueva de criatura viva, una mutación que, a diferencia de la mayoría de las mutaciones, era apta para sobrevivir. El conjunto de formas félidas posibles había rebosado hasta invadir un nuevo nicho ambiental, una nueva colina del espacio de competencia, y se había establecido en él. ¡Las mutaciones aleatorias, actuando a modo de "ruido", habían expulsado a una sección de la especie gatuna de uno de los máximos locales, y la habían impulsado hasta otro máximo local!



6. Efecto de la mutación aleatoria en las criaturas que ocupan un nicho ambiental. (a) *Un conjunto de formas posibles* (borrón verde) *confinadas en un nicho.* (b) *Los cambios aleatorios provocan que el conjunto proyecte un sensor hacia un nicho vacante vecino.* (c) *Parte del conjunto se escinde y forma una especie nueva.*

¿Explicaba esto los súbitos cambios en el registro fósil?

La verdad es que no.

En primer lugar, las mutaciones del tipo que acababa de observar serían sumamente insólitas. Esta es, de hecho, una de las objeciones típicas que se le oponen a la teoría darwinista. La serie completa de cambios de forma necesarios para crear de un solo golpe una especie verdaderamente nueva tendría que ser muy rara. Quizá yo había tenido suerte, o había elevado demasiado la tasa de mutación.

Además, cuando en el registro fósil se produce un cambio súbito de especies, lo que suele ocurrir es que aparezcan al mismo tiempo *varias* especies nuevas. En lugar de bifurcaciones (escisiones en dos) se producen *multifurcaciones* (nacimiento simultáneo de varias ramas). Ya Darwin se había percatado de este hecho; tanto así, que el único diagrama que hay en todo *El origen de las especies* tiene la función de ilustrar este preciso punto. En cambio, si la evolución de una especie nueva se produjera debido a que sus criaturas representativas "se desbordasen" y "salpicasen" al azar a unos nichos desde otros, solamente debería producirse una especie nueva en cada ocasión.

Por último, en el mecanismo anterior, la especie primitiva subsiste: ninguna especie llega a extinguirse; no resulta muy plausible. Cabe, claro está, la posibilidad de que todo el grupo de gatos pasase a ocupar el nicho de los gatos alados; pero eso es mucho menos verosímil que el que lo haga sólo una reducida parte de ellos. Tampoco resulta muy plausible.

Lo que yo deseaba hallar en este modelo matemático era un método más robusto para la evolución de nuevas especies. Un párrafo prometedor del manual parecía permitirlo.

Variación ambiental. En el mundo real, el nivel de viabilidad de una criatura depende de su ambiente. Por ejemplo, una criatura que por evolución ha llegado a poseer un largo cuello para poder comer las hojas de árboles altos perderá su ventaja adaptativa y resultará menos apta para sobrevivir si la sequía reduce la altura de los árboles. El mando de *variación ambiental* permite al usuario experimentar con distintos tipos de cambios, lo mismo en tiempo que en espacio.

Se sugiere que, en el primer ensayo

de cambio ambiental, la tasa de mutación al azar se ajuste a cero o, en todo caso, a un valor muy pequeño.

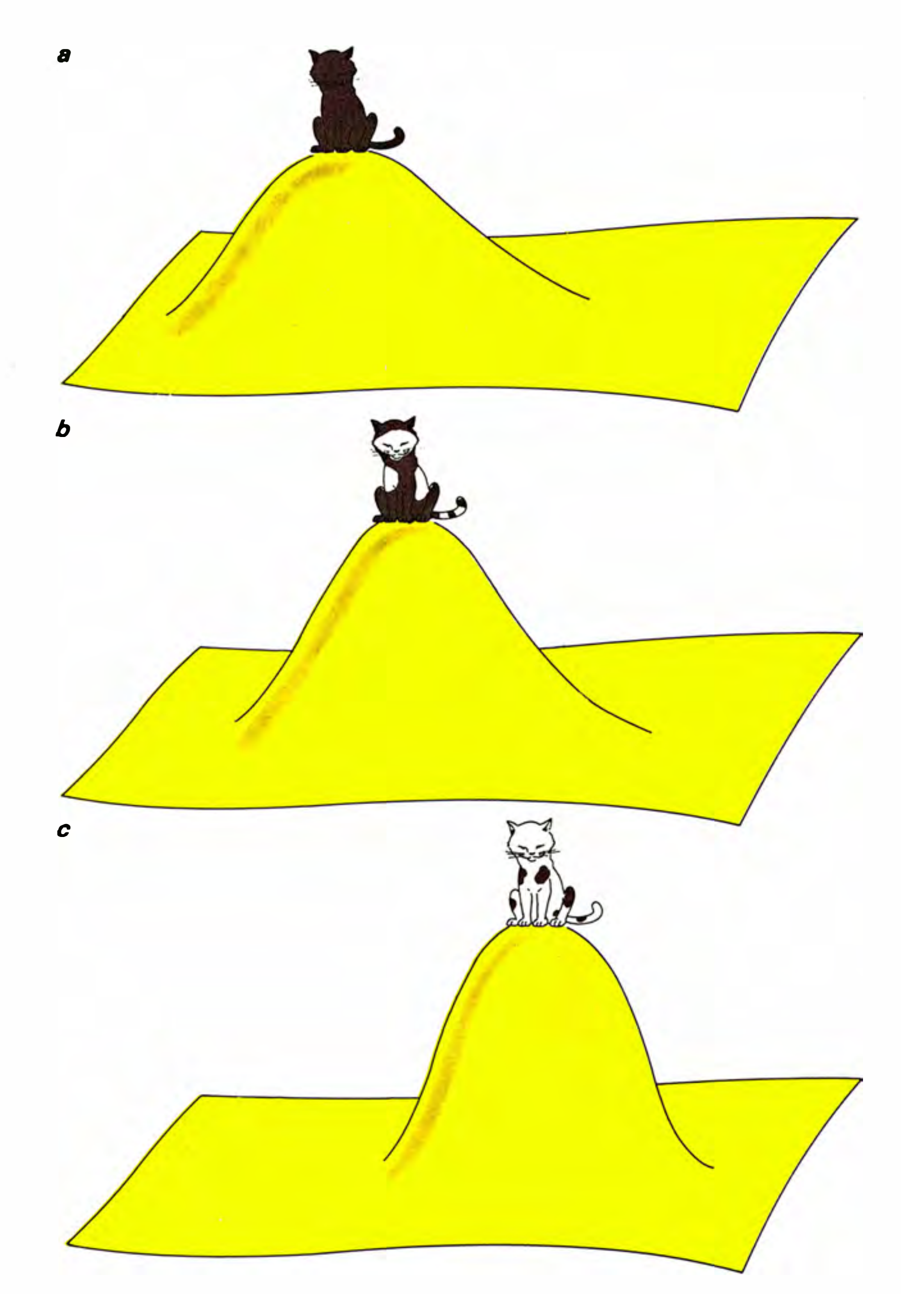
Perfectamente, pensé. *Probemos con la variación temporal*. (Aunque la variación espacial también es interesante, dejaré que sea el lector quien experimente con ella). Hice avanzar un poco el mando.

El paisaje comenzó a ondular, como un mar tormentoso visto a cámara lenta. Los grupos de criaturas viables se aferraron a sus nichos ambientales, encaramándose en la cima de su altillo particular al ir éste abombándose más, contrayéndose, subiendo, bajando, o desplazándose hacia un lado como un todo [*véase la figura 7*]. La especie era capaz de responder con suficiente rapidez y cambiar adaptativamente, manteniendo máximas su competitividad y aptitud para la supervivencia. Comprendí que estaba presenciando la *estabilidad intrínseca* de una especie particular: en un ambiente que varía con lentitud, también la forma de una especie cambia lentamente, si es que llega a hacerlo. He aquí pruebas en favor de una teoría gradualista, pensé. Pero, ¿qué ha sido eso?

Con el rabillo del ojo vi una perturbación repentina, violenta. Continué observando cómo el paisaje se rizaba, como ocurre al arrojar una piedra en un estanque; fue entonces cuando lo vi. Una especie entera salió súbitamente disparada ladera arriba de un montecillo cercano, cambiando de forma en un instante al ocupar un nuevo nicho.

¿A qué venía ese traslado desde su antiguo nicho? Y ya que lo menciono... ¿dónde estaba su antiguo nicho?

Si esto fuera un programa de televisión, he aquí una ocasión excelente para repasarlo a cámara lenta. Podríamos ver entonces lo realmente sucedido... un momento, ¿para qué sirve ese botón?



7. Al ir un nicho moviéndose gradualmente o ir cambiando de forma, la especie que lo ocupa va también cambiando gradualmente.

Estaba rotulado *repetición lenta*. Muy bien. Lo pulsé.

La serie de acontecimientos resultaba obvia en este caso [*figura 8*]. El desplazamiento rápido se producía al *desaparecer* un nicho ambiental. Al principio no había más que un nicho: un cerrillo de poca altura, ocupado. Después, en la vecindad, empezó a crecer otro, desocupado, donde antes no había más que una región indiferente; la variación ambiental había creado el potencial de una nueva especie. Pero el nuevo nicho estaba vacante, porque ninguna criatura había acertado a efectuar la transición necesaria para ocuparla. El nuevo nicho creció, y pronto fue más alto que el antiguo. Se desplazó hacia el viejo. Entre ellos se encontraba un puertecillo, un punto de ensilladura de la superficie... El puerto se elevó hasta la altura del antiguo nicho; y los dos se cancelaron mutuamente, como una partícula al hacer impacto con su antipartícula. Pero en vez de liberar energía, la colisión liberó a las criaturas de la trampa que su antiguo nicho suponía. Libres ahora para mejorar su valor adaptativo, salieron disparadas pendiente arriba de la nueva colina, zigzaguearon un poco de acá para allá, y finalmente ocuparon la cumbre. Pero ahora su aspecto era diferente: habían evolucionado para adaptarse a un nicho nuevo.

¡Y a qué velocidad! Parecía que esta vez los puntualistas iban a anotarse un punto a favor.

¡Tal vez las dos teorías fueran correctas! Me daba la impresión de que había dos fenómenos distintos. Ambos tenían una misma causa general, la variación ambiental, pero se producían en circunstancias diferentes.

Variación continua y lenta: el nicho ambiental cambia de forma o se desplaza, pero sigue siendo el "mismo".

Cambio súbito, efectivamente discontinuo: Un nicho ambiental entra en colisión con otro y desaparece.

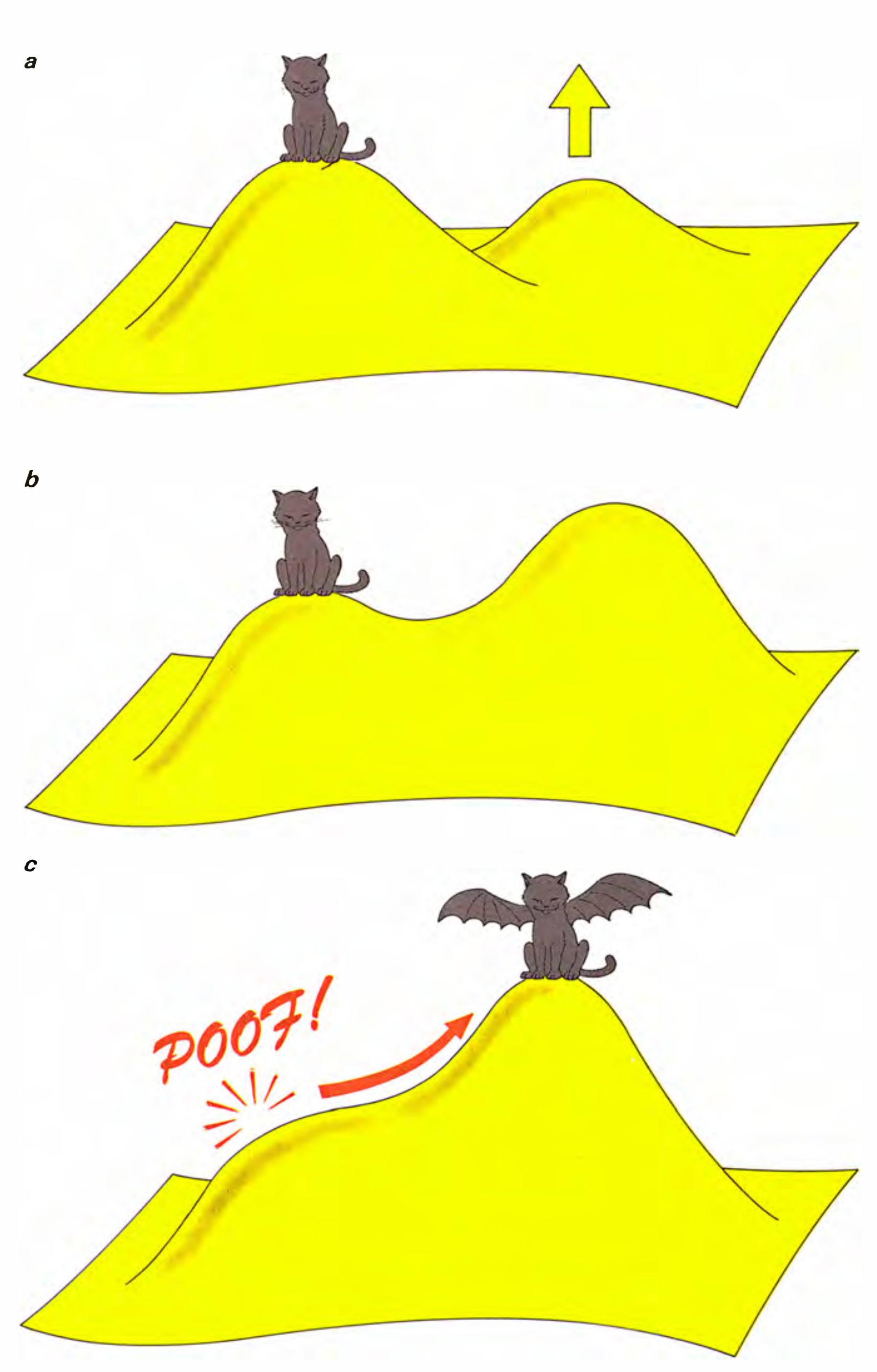
La característica a destacar es que tanto el cambio gradual como el equilibrio puntual resultan ser consecuencias naturales del mismo modelo matemático de la evolución darwinista.

Además, aun cuando el modelo matemático no respondiera a la auténtica evolución, todavía mostraría que es posible la coexistencia de ambos tipos de cambio, sin causas especiales y distintas, y que no existe *necesidad lógica* de tenerlas por teorías en conflicto. Dicho de otro modo, quizá gradualistas y puntualistas debieran sumar sus mentes en lugar de darse cabezadas unos contra otros.

El libro tenía una nota a pie de página:

La teoría matemática de *singularidades*, también conocida por *teoría de catástrofes*, inventada por René Thom, Vladimir Arnold, John Mather y otros, muestra que los dos tipos de comportamiento antes observados son los fenómenos típicos —y los únicos fenómenos típicos— que pueden tener lugar en un modelo matemático de esta clase. La diferencia está en que el cambio lento tiene lugar durante casi todo el tiempo, mientras que los cambios súbitos (o *catástrofes*) sólo acontecen en instantes aislados. Pero esto es precisamente lo que ocurre en el registro fósil. Esta aplicación de la teoría de catástrofes a la evolución se debe a Christoffer Zeeman y es continuación del trabajo de Maurice Dodson y otros.

Curiosamente, los primitivos proponentes de los cambios bruscos en el proceso evolutivo, allá por los tiempos de Darwin, fueron calificados de *catastrofistas*.



8. Evolución "catastrófica" de una nueva especie, provocada por la desaparición de un nicho.

En este punto, caí en la cuenta de que no todas las dificultades habían quedado solucionadas. El tipo de cambio súbito que tiene lugar en el modelo —y según la teoría de catástrofes, el único tipo posible de salto, excepto en casos sumamente excepcionales— conlleva la *desaparición* de una sola especie y la creación de, así mismo, una sola especie nueva. Lo cual no explica la observación de Darwin de la creación de varias especies al mismo tiempo, emanadas, todas ellas, de una misma especie precursora. Al cabo de un rato de desconcierto, me vino a la memoria que había puesto a cero el mando de *mutación aleatoria*. Quizás ello crease alguna diferencia, aunque no me imaginaba cómo.

Repetí el experimento, esta vez con las mutaciones aleatorias activadas, con lo que en lugar de tener una única especie encaramada en lo alto de cada nicho, había un borrón vibrante de formas disímiles pero emparentadas, que ocasionalmente descendían un poco ladera abajo, aunque, por lo común, el glóbulo permanecía confinado en torno a la cima por selección de los más aptos. El primer ensayo no produjo apenas diferencia alguna: en cuanto el nicho desapareció, el borrón entero ascendió rápidamente por la nueva colina y acabó en el mismo lugar que antes.

Pero observé que, al subir por la ladera, el glóbulo tendía a abrirse hacia los lados. Existe una buena razón para que así sea. Pensemos en una forma particular de criatura, rodeada por un borrón circular de variaciones aleatorias de dicha forma. Las formas dotadas del mismo nivel de adecuación se hallan situadas en un mismo isocontorno (curva de nivel) del paisaje competitivo. Son, por consiguiente, igualmente aptos, y tienen pues la misma tasa de supervivencia. Las formas situadas ladera abajo son menos aptas y, por eso mismo, menos capaces de sobrevivir. Así, pues, aunque el borrón circular se encuentra comprimido desde atrás, puede desparramarse hacia los lados; tiende a

parecerse a una elipse, achatada a lo largo del borde inferior.

¿Y si un borrón de formas posibles ascendiera por una pendiente situada en la vecindad, no de un cerro, sino de varios? Al abrirse hacia los lados, puede fragmentarse en piezas que ocupen *todas* las colinas vecinas [*véase la figura 9*]. Lo cual conduce a especiación múltiple, como había observado Darwin.

El "árbol evolutivo" correspondiente ofrece el aspecto de la *figura 10a*; pero, dado que los sedimentos litogénicos se van depositando con mucha lentitud, la escala temporal queda eficazmente comprimida y lo que vemos en el registro fósil se asemeja más a la *figura 10b*, que recuerda a las multifurcaciones de la figura 1. ¡Todo encaja!

Así que, cuando acoplamos los dos ingredientes esenciales de la teoría de Darwin, a saber, mutación aleatoria *más* supervivencia de los más aptos—, el modelo matemático más natural y simple posee las siguientes propiedades:

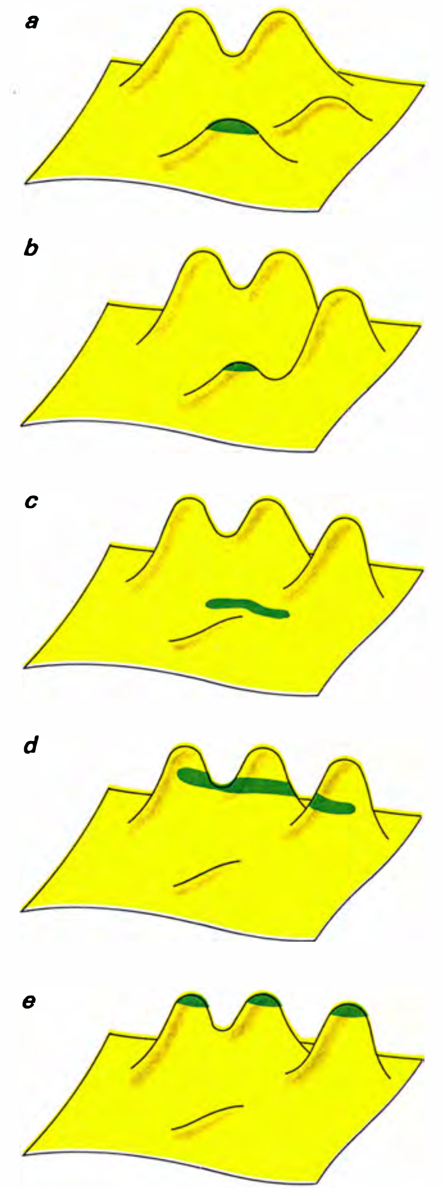
—Durante casi todo el tiempo, las especies cambian lentamente, o no cambian en absoluto.

—Durante los cambios lentos, no hay creación de nuevas especies.

—En ocasiones aisladas, las especies pueden cambiar mucho más rápidamente.

—Cuando se producen cambios catastróficos súbitos, es frecuente la evolución simultánea de especies múltiples.

Fijémonos en que las catástrofes no son meras versiones aceleradas de los cambios lentos. Cualitativamente, son del todo diferentes. Matemáticamente, comportan la desaparición de un nicho, mientras que los cambios lentos sólo comportan la modificación continua de un nicho. Además, ni siquiera una versión acelerada de un cambio "lento" puede producir una especie nueva; solamente una catástrofe puede. Todo esto se aproxima tanto a lo observado en las series fósiles que arroja dudas sobre si el gradualismo y el puntualismo constituyen en realidad teorías en conflicto.¡Tal vez no sean más que facetas distintas de la *misma* teoría! ¡Tal vez ni el gradualismo, ni el puntualismo estén en lo cierto, por separado, sino que lo estén ambos!



9. "Multifurcación" de especies causada por la desaparición de un nicho en presencia de variaciones aleatorias. La mancha verde representa un conjunto de seres. (a) *Una especie individual con varios nichos desocupados próximos.* (b) *El nicho ocupado comienza a ser absorbido por los vecinos vacantes.* (c) *Al apartarse el conjunto del equilibrio evolutivo, se extiende lateralmente a lo largo de las líneas de adaptación constante.* (d) *El conjunto invade tres nuevos nichos.* (e) *La formación simultánea de tres especies nuevas, de nuevo en equilibrio evolutivo.*

Evidentemente, el modelo de que hablamos no es sino el más sencillo. Existen otras muchas formas, aparte de la teoría de catástrofes, de efectuar una traducción matemática de las hipótesis de Darwin. Otros modelos, más complicados y realistas, se ajustan aún mejor a los registros fósiles. Pueden combinar los cambios rápidos y lentos en una teoría coherente. De hecho, el relativo éxito de la sencilla idea aquí esbozada sugiere que así debería ser, que esta combinación es consecuencia natural de principios matemáticos generales, y no un hecho arbitrario.



10. (a) *Arbol evolutivo correspondiente a una multifurcación, como la de la figura 10.* (b) *El mismo árbol, en el registro fósil, ofrece un aspecto sumamente comprimido, lo que produce aparentes discontinuidades. Compárese con la figura 1.*

El manual concluía con una nota de advertencia:

¡AVISO! El evoluscopio es un prototipo. En modelos más perfeccionado se incorporarán muchas características no presentes en él. Entre ellas,

Refinadas medidas multivariantes del grado de adecuación, en lugar de una única medida numérica, escalar.

—Retroalimentación desde la competencia entre individuos al ambiente, para modelizar la posibilidad de que los cambios en los competidores alteren las reglas de la competencia.

Este segundo punto era excelente. ¡La evolución de un gato alado podría tener efectos dramáticos en el paisaje competitivo de los pájaros!

—Atención debida a la diferencia entre *genotipo* (el material genético de la criatura) y el *fenotipo* (su forma). La selección natural actúa sobre el fenotipo, mientras que las mutaciones aleatorias actúan sobre el genotipo. Por consiguiente, el modelo ha de dar cabida a la interacción entre dichos factores a través del proceso reproductivo...

Libros

Ecología teórica, historia de la física, álgebra fundamental y ciencia egipcia

R. Margalef, M. García Doncel, E. Elizalde y L. Alonso

PERSPECTIVES IN ECOLOGICAL THEORY. Dirigido por Jonathan Roughgarden, Robert M. May y Simon A. Levin. Princeton University Press; Princeton, 1989.

Este volumen es resultado de un encuentro que en 1987 reunió unas 50 personas en Asilomar. Entre ellas figuran los nombres que más se han distinguido en el cultivo de la teoría ecológica. Es, por tanto, un buen documento para hacerse cargo del estado actual de esta disciplina. Quizás esta presentación se complace demasiado en temas repetidos ya *ad nauseam*, por lo menos en ciertas escuelas, y se preocupa menos en explorar o en señalar otras direcciones que ahora son posibles fuera de los caminos más trillados. Tal renovación resulta urgente, ante el desarrollo de una ecología planetaria, que pide conceptualizar mejor el engarce entre la biosfera y el entorno físico. En este volumen, después de una introducción redactada por los editores, el material aportado por la mayoría de los participantes a la reunión se ha distribuido en 8 secciones, cada una de ellas con uno o dos trabajos, generalmente en forma de revisiones, más otra contribución orientada hacia la discusión crítica. Además de las referencias específicas al fin de cada artículo, se incluye una lista de libros más generales. El formato general es apropiado, la tipografía es agradable y las figuras son claras. Los siguientes comentarios destacan algunos aspectos que han llamado más la atención del que suscribe este comentario, o que conciernen en grado diverso a sus manías personales, pero, por supuesto, el contenido es mucho más rico, y admite otras lecturas. Sin embargo, estas líneas pueden dar idea del interés de la obra.

El título de la primera sección es "de los individuos a las poblaciones" y trata, por una parte, de la fisiología vegetal, de la fotosíntesis y de la arquitectura de las plantas superiores, con la habitual preocupación por la economía del agua, pero con menos sensibilidad hacia el reconocimiento del significado del sistema de sostén y transporte. Se continúa con el estudio del comportamiento y distribución de los organismos en función de la distribución de los "recursos". Parte del debate se centra en la alternativa entre una explicación mecanística o una descripción fenomenológica.

La segunda parte trata de dinámica de poblaciones e interacciones entre especies, tema que se aborda con talante optimista. Contiene un largo artículo, con buena bibliografía, de Pacala, sobre la dinámica de poblaciones en plantas superiores. Sigue, en general, la interpretación que da Tilman a los modelos de Lotka-Volterra y contempla su aplicación a una teselación de celdillas sobre un plano; este enfoque conduce a reconocer la caída de la diversidad después de la fertilización y los aspectos esenciales de la sucesión. Como es, por desgracia, habitual, se utiliza poco la experiencia de los forestales. P. Kareiva, trata de las relaciones entre teoría y experimentación en ecología. Yo tuve buena relación y amistad personal con MacArthur, que tenía una fina percepción de la naturaleza, y no comparto la opinión de que MacArthur relegara la dinámica de las poblaciones, en favor de la estática, y que diera poca importancia a los experimentos en botellas —en todo caso los ponía en su lugar—. Es cierto que hacen falta más experimentos en pequeños recintos, a lo que se puede añadir que ya sería hora de imitar a Huffaker (1958!) y no solo alabarlo. Uno no puede menos que reconocer la razón que lleva Kareiva al decir que "la triste verdad es que la teoría ecológica existe, en buena parte, en un mundo que le es propio, ignorado por los avances de la ecología". Pero esto en parte es culpa de la poca sustancia de la teoría ecológica de que actualmente disfrutamos.

La sección 3, que trata de ecología y evolución, procura unificar la biología de las poblaciones, incorporando la genética. Los experimentos de selección suelen afectar a muy pocos genes, lo que plantea dificultades a la hora de interpretar la evolución lenta en el seno de los ecosistemas, especialmente la evolución de las llamadas estrategias vitales. Es cierto que hay muchas dificultades en la representación matemática de las teorías de expresión verbal y prolija, lo cual se manifiesta en relación con la dependencia de la densidad, y de las diversas y criticables estrategias asociadas tradicionalmente a los símbolos r y K. La evolución de la senescencia es un tema importante, después de constatar la tendencia en las estirpes a alargar la edad individual y el posible desarrollo de una tolerancia a una mayor densidad de población. La interpretación ecológica del registro fósil debería ofrecernos la necesaria perspectiva a gran escala. Darwin veía un mundo muy competitivo y trabado, donde los cambios evolutivos importantes debían hacerse difíciles. Al modelo gradual tradicional, se opone la evolución más rápida, que ahora se llama puntuada. De una manera u otra ha estado siempre presente en el ánimo del paleontólogo, aunque nunca se ha explicado adecuadamente. Tampoco existen modelos sencillos y convincentes de la coevolución.

El tema de la escala y acoplamiento en sistemas ecológicos (parte 4), incluyendo consideraciones sobre la jerarquía (R. V. O'Neill), es, a juicio del que suscribe, el que tiene más aire de novedad, tal vez por referirse principalmente a sistemas acuáticos (T. M. Powell, J. H. Steele), que pueden ser más fáciles de entender y formular. Sería bueno que los estudiosos de ecosistemas terrestres se inspiraran más en lo que ofrecen los acuáticos, hasta llegar a una formulación común de algunos problemas y conceptos de ecología planetaria, de muestreo a distintas escalas y de cómo persisten distribuciones en mosaico (cuencas) o en forma de serie lineal más o menos perturbada (ríos). La jerarquía depende de muchas cosas, por ejemplo, de la relación entre la extensión del sistema y la tasa de renovación global —y de la relación entre el tamaño de los organismos y el tiempo de duplicación de sus respectivas poblaciones—; la evolución puede ir más rápida a pequeña escala; a mayor escala se convierte en un paseo al azar. Lo cual no favorecería precisamente a los macrófitos, a pesar de que han tenido indudable éxito en otros aspectos. Semejantes consideraciones en relación con la jerarquía resuelven los falsos problemas sobre si el control en las redes tróficas opera de arriba abajo, o de abajo arriba. El resultado de la competencia, la probabilidad de éxito, tiene que ver con la escala. Se reconoce la necesidad de estudiar los

motivos de distribución en el espacio y en el tiempo. (Así se ve que es imposible estabilizar las variancias, que es una de las críticas más serias que puede hacerse a la ciega aplicación de una estadística elemental a la ecología).

Powell discute, con referencia al mar, la noción de escala, la distancia a recorrer para que cambie significativamente alguna magnitud de interés. En el mar nos movemos entre la viscosidad a pequeña escala y el efecto de Coriolis a gran escala; entre procesos biológicos a pequeña escala y los fenómenos a la escala de El Niño, relacionados con un acoplamiento, a escala planetaria, de la atmósfera, los océanos y aun los continentes. El acoplamiento entre los regímenes diversos no es linear. Según Powell, al aumentar la escala, aumenta la variancia, y así parece bien claro en sistemas oceánicos. Sin embargo, según Levin (sección 6), en sistemas relativamente cerrados, se observa una relativa disminución de la variancia al aumentar la escala. De todos modos, la variancia no es estable, y las situaciones posibles llevan a pensar al lector que los ecosistemas son muy diversos, pudiendo ser calificados de (localmente) "convexos" o "cóncavos".

La sección 5 trata de la estructura y ensamblado de las comunidades. J. E. Cohen se ocupa de las cadenas tróficas, en cuya interpretación separa la visión fisicoquímica de la naturalística, lo cual quizá no es muy apropiado, porque la historia natural es la base para reconocer la asimetría formal de cada una de las relaciones tróficas. Prosigue la presentación, usando "especies tróficas", un concepto que en la práctica se aplica a cualquier conjunto arbitrario, y "redes tróficas" tal como han sido publicadas. Su estudio comparado conduce a formular relaciones, ya dadas a conocer anteriormente, que forman parte de un sistema de alometrías ecológicas. Roughgarden se plantea, una vez más, el problema de si las comunidades están formadas por especies que han podido llegar hasta el punto que se considera y cuáles son las restricciones que las posibilidades de su coexistencia plantean. Se comenta brevemente otra venerable reliquia, el tema de las relaciones entre diversidad y estabilidad. En la última década se ha hecho un esfuerzo considerable (unos 200 estudios) para definir la importancia de la competencia, en las condiciones en que naturalmente se ensamblan las comunidades, condiciones que pueden ser complicadas. Los modelos de Lotka-Volterra deben tratarse como metáforas, ayudan poco a entender y menos a resolver problemas concretos.

La estructura y función de los ecosistemas es el tema de la sección 6. S. A. Levin destaca el interés de la simulación, y la importancia del espacio. Sus simulaciones conducen a resultados, ya comentados, que complementan los de Powell. En la discusión general se repiten las habituales jaculatorias.

Las secciones 7 y 8 se ocupan principalmente de temas prácticos, la 7 de ecología y manejo de recursos y la 8 de ecología de plagas y patógenos. La sección 7 produce la impresión de *dejà vu*, considerando, entre otros temas, la fabulosa curva que relaciona el esfuerzo de pesca con el rendimiento. Entre las causas de la amenaza de una pérdida catastrófica de diversidad biológica se cuentan la destrucción y fragmentación del habitat, las matanzas por el hombre y la introducción de especies exóticas, así como consecuencias de la extinción de otras especies. Sin embargo, no se habla de las posibles consecuencias de la aceleración de la dinámica de los ecosistemas, bajo la explotación humana. Paul Ehrlich se pregunta hasta qué punto la teoría ecológica ha sido útil en la práctica, haciendo sensatas consideraciones que llevan a pensar en lo muy conveniente que es buscar una más estrecha relación con la economía. La última sección contiene un interesante artículo de Hassell y May sobre modelos de las asociaciones entre parásitos (o parasitoides) y sus hospedadores, partiendo de los modelos tradicionales, pero introduciendo diversas matizaciones (heterogeneidad espacial, dinámica no lineal, mecanismos dependientes de la densidad) que prometen un mayor realismo.

En resumen, este libro ofrece una visión panorámica de la relativa confusión, casi desesperación, que aqueja a la ecología teórica actual y de los temas que ahora se llevan más. En este sentido es muy útil, aunque, por otra parte, resulta decepcionante, por la escasez de cosas nuevas y aprovechables. El autor de esta reseña publicó en 1968 un librito (111 páginas) con el mismo título que este volumen y ésta fue una de las razones que me movieron a leer el presente con considerable avidez. Pero, o bien el horizonte queda tan lejano que no podemos apreciar si nos movemos, o bien la teoría ecológica hace ya más de un cuarto de siglo que está en período de *stasis*. (R. M.)

EL NACIMIENTO DE UNA NUEVA FÍSICA, por I. Bernard Cohen. Alianza Universidad; Madrid, 1989.

Bernard Cohen es bien conocido como editor crítico de los *Principia* de Newton, y consiguientemente como uno de los más prestigiosos investigadores de la evolución conceptual newtoniana. La presente obra es un valioso trabajo de divulgación sobre la revolución científica coronada por la obra de Newton. Concebido en 1960, en el marco de una colección y una comisión que pretendían un nuevo enfoque didáctico para la comprensión de las ciencias y especialmente de la física, fue dedicado a su hija. Al reeditarlo veinticinco años más tarde, Cohen ha enriquecido la parte relativa a Galileo, esforzándose en hacer llegar al nivel de la divulgación el fruto obtenido mediante la investigación de nuevos manuscritos galileanos.

En su forma actual, la obra está dividida en siete capítulos y dieciséis apéndices. El primer capítulo sirve de introducción, y sugiere la tesis global del libro: la concepción copernicana de una tierra en movimiento exige una nueva física que será elaborada por Galileo, Kepler y Newton. El segundo divulga la física aristotélica y en especial su ley de velocidad de caída (proporcional directamente a la gravedad e inversamente a la resistencia). El tercero presenta las novedades y los problemas del sistema copernicano. Los capítulos cuarto y quinto dan a conocer la figura de Galileo como observador astronómico y como físico teórico y experimentador. El sexto describe las leyes de Kepler y su contexto de búsqueda de las armonías celestes. Por fin, el capítulo séptimo presenta como coronación la nueva física de Newton, a la vez celeste y terrestre.

Sus apéndices son muy breves, algunos ocupan una página escasa. Los diez primeros tratan diversos aspectos de las investigaciones galileanas: su telescopio, sus observaciones astronómicas, sus experimentos, su fundamentación experimental, sus fallos iniciales sobre la ley de caída libre, su método hipotético-deductivo, su relación con el mundo medieval, su concepto de inercia comparado con los de Kepler, Descartes y Gassendi, su descubrimiento de la trayectoria parabólica y, finalmente, un resumen de sus aportaciones más importantes. Los seis últimos apéndices abordan diversos aspectos de la investigación newtoniana: su deuda con Hooke, su visión de la inercia en el movimiento planetario, su encuentro temprano de la ley cuadrático-inversa, su encuentro tardío de la contrastación lunar, su concepto de masa inercial y gravitatoria, y sus últimos pasos hasta llegar a la idea de gravitación universal.

No se busquen en este libro novedades históricas, pero sí una divulgación fácilmente asequible a nivel de bachillerato de lo que nos ha aportado la investigación histórica, la suya propia en torno a Newton y la de otros investigadores en torno a Galileo. Entre éstos, cita continuamente a Stillman Drake. Cierto que recorta a veces sus exageraciones galileolátricas, sea aportando las tesis contrarias de otros investigadores (W. L. Wisan y R. H. Naylor, p. 203), sea aportando las suyas propias (p. 212). No obstante, a nuestro juicio, Cohen exagera un poco el valor de la física galileana, más allá de sus indiscutibles logros cinemáticos. Subraya repetidas veces una aportación dinámica inicial (pp. 97, 124, 131, 186), y el que su concepto de inercia no puede llamarse una inercia circular (p. 126; por más que al presentar el texto galileano aparezca como tal: p. 129), minimiza como teorizante toda especulación de los precursores (pp. 112, 207, 215, 216), y le atribuye a Galileo un concepto de aceleración, como magnitud física, que ciertamente Galileo no tenía (pp. 101, 106, 107, 109, 111, 131, 201). Se echa de menos, por el contrario, la exposición del principio galileano de los planos inclinados (p. 105), que no menciona hasta el resumen final (p. 222), y el mérito teórico que supone la primitiva integración galileana de la velocidad en función del tiempo. Finalmente expone, y aun ilustra con una figura (fig. 21, p. 119), una experiencia que resulta muy didáctica, pero que no es fácil ver relacionada con los textos galileanos, como da a entender allí mismo con sus citas (pp. 219-220). Quizá se trate de un sencillo juguete moderno de inspiración galileana, como la simpática locomotora eléctrica que ilustra fotográficamente lanzando verticalmente por la chimenea una bola que cae inexorablemente sobre la misma chimenea, tanto si la locomotora está parada como si avanza con velocidad uniforme. Es un juguete de indiscutible valor didáctico, para intuir desde niño ese principio de inercia que tanto costó descubrir a la humanidad.

El rigor del capítulo newtoniano es indiscutible, aunque quizá resulte algo denso. Queda, sin embargo, muy bien ilustrado con el conjunto de apéndices newtonianos que gozan de una gran vitalidad, dentro de su concisión.

Prescindiendo de detalles nimios, la obra en su conjunto es muy seria. Revela ese raro mérito de conjugar en un mismo autor un verdadero esfuerzo de divulgación de los conceptos científicos con un conocimiento de investigador profesional sobre el proceso de su evolución histórica. No podemos decir lo mismo de la labor del traductor. A veces hace el texto incomprensible, imaginando por ejemplo un libro de Osiander (p. 91), o un misterioso movimiento hacia arriba "por" la circunferencia "a la fuerza" (p. 127). Otras veces simplemente lo dificulta con su terminología: movimiento "complejo" por "compuesto", curva "oviforme" por "oval", proposición "conversa" por "recíproca", "prueba de la luna" por "contrastación lunar". (M.G.D.)

ALGEBRA I: BASIC NOTIONS OF ALGEBRA, por I. R. Shafarevich. Springer Verlag; Berlín, 1990.

Este texto forma parte de la "Enciclopedia de las Ciencias Matemáticas", obra ambiciosa que recopila toda una serie de artículos de autores soviéticos, bajo la supervisión general de R.V. Gamkrelidze. La traducción inglesa la está editando Springer Verlag; de ella han aparecido, además del 11 que aquí se reseña, los volúmenes 1 al 5, que tratan de la teoría de los sistemas dinámicos (I a V); el 7 y 9, sobre varias variables complejas (I y III), y el 13, titulado *Análisis I*. Si todos ellos mantienen el mismo nivel de calidad que el presente, será ésta una enciclopedia que no podrá faltar en la casa de ningún matemático, físico teórico o estudioso de las teorías científicas en general.

Según confiesa el autor, su objetivo es presentar una visión general del álgebra, de sus nociones básicas y de sus múltiples ramas. En lugar de elegir para ello el camino (tantas veces utilizado por desgracia) de la enumeración detallada de miríadas de estructuras y de las relaciones que existen entre las mismas, Shafarevich ha procurado extraer con buen criterio la pequeña porción de ellas de interés real para los matemáticos y para los científicos teóricos. Además, jamás se ha limitado a la mera definición formal de la estructura, sino que se ha preocupado de ilustrarla (diría incluso de expresarla, en pie de igualdad con la definición rigurosa) con unos cuantos ejemplos básicos que motivan la propia definición y nos iluminan sobre su significado preciso. No sé si esto será apreciado en todo su valor por los matemáticos puros ni por los especialistas en el tema, a quienes todas estas ejemplificaciones puedan parecer superfluas. En mi caso he de confesar el profundo asombro y satisfacción que experimenté ya con el ejemplo inicial del capítulo primero: "¿Qué es álgebra?".

Consta el libro de 22 capítulos. Del primero ya he hablado. El segundo trata de los axiomas de cuerpo, con ejemplos como el de funciones racionales de varias variables y el de las series de Laurent formales, entre otros. El capítulo 3 nos introduce en los anillos conmutativos, cuerpo de fracciones, anillos polinomiales, booleanos, etc., así como en la factorización. El 4 se dedica a homomorfismos e ideales, anillos cociente y representación de anillos; se habla también en él de los ultraproductos y del análisis no estándar. El capítulo 5 está consagrado por entero a los módulos, con todas sus operaciones básicas y con ejemplos de módulos de formas diferenciales y de campos vectoriales. El capítulo 6 aborda los aspectos algebraicos de la dimensión. Tienen en él cabida los módulos de tipo finito, los módulos y anillos noetherianos y las extensiones finitas. El capítulo 7 ofrece la visión algebraica de las nociones infinitesimales, como la de espacio tangente de una variedad, campos vectoriales y operadores diferenciales, de los infinitésimos de orden superior, de los "jets" y de los campos de números p-ádicos en teoría de números.

El capítulo 8, sobre los anillos no conmutativos, recoge entre otros conceptos el de anillo de endomorfismos de un módulo, el de álgebra grupo, las álgebras de división, las fibraciones de "twistors", el álgebra exterior, el álgebra de Clifford y las superálgebras. En el capítulo 9 se explican los módulos sobre anillos no conmutativos, representaciones matriciales, el teorema de Jordan-Hölder, el lema de Schur, etc. En el 10 se introducen los módulos y anillos semisimples, con el teorema de Wedderburn, las geometrías continuas y aplicaciones a la representación de grupos finitos. Sigue, en el 11, con las álgebras de división de rango finito; en él se consideran, en particular, el teorema de Tsen y los cuerpos casi algebraicamente cerrados, así como las álgebras de división sobre cuerpos p-ádicos y de los racionales.

La noción de grupo se introduce en el capítulo 12 y a esta estructura fundamental del álgebra se aplican siete capítulos del libro. Choca esta ordenación con la que estamos acostumbrados a encontrar en muchos textos al uso. Evidentemente, la estructura de grupo es la más simple en su definición, la menos constreñida en sus propiedades. De ahí que resulte la más complicada en su estudio, lo que explica la extensión y la situación que Shafarevich le adjudica en su obra. En los capítulos 13, 14 y 15 se pasa revista a los ejemplos más importantes de grupos. En el 13, en particular, aborda los grupos finitos: simétricos y alternantes, los de retículos, las clases

cristalográficas y los generados por reflexiones. En el 14 se estudian los grupos discretos infinitos, grupo modular incluido, los grupos libres y los de nudos y trenzas. El capítulo 15 trata de los grupos de Lie —en particular los de Lie compactos y los de Lie analíticos complejos— y de los grupos algebraicos.

El capítulo 16 se dedica a una exposición de los resultados generales de la teoría de grupos; termina con la clasificación de los grupos finitos simples. En el capítulo 17 se estudian las representaciones de grupos, en especial, las representaciones de grupos finitos, las de grupos de Lie compactos y las representaciones de los grupos de Lie complejos clásicos. En el capítulo 18 se tratan algunas aplicaciones de los grupos, como son la teoría de Galois, la teoría de Picard-Vessiot, o teoría de Galois de las ecuaciones diferenciales lineales, la clasificación de los recubridores no ramificados y el grupo fundamental, el primer teorema fundamental de la teoría de invariantes; acaba con la aplicación de las representaciones de grupos a la clasificación de las partículas elementales.

El capítulo 19 expone álgebras de Lie y otras álgebras no asociativas importantes, como los números de Cayley, algunas estructuras casi complejas y álgebras de división no asociativas. En el capítulo 20 se introducen las categorías. En el 21, el álgebra homológica, con sus orígenes topológicos, las cohomologías de módulos y de grupos, las cohomologías de haces y el teorema de Riemann-Roch. Finalmente, el capítulo 22 se dedica a la teoría K en sus dos aspectos: la teoría K topológica y la teoría K algebraica. Termina el texto con unas páginas de comentarios sobre la bibliografía, que abarca 115 títulos.

No es pretensión del autor con este libro enseñar álgebra, sino únicamente "hablar sobre ella". El defecto lo subsana con la extensa bibliografía del final de la misma. El libro está dirigido, en palabras del propio autor, a estudiantes de matemáticas de los primeros cursos universitarios, así como también a físicos teóricos y a matemáticos no especialistas en álgebra. Como representante del segundo y del tercer grupo en tal clasificación de interesados en esta disciplina apasionante cual es el álgebra, repito mi opinión de que el autor ha conseguido plenamente su objetivo y, con él, los responsables de la "Enciclopedia de las Ciencias Matemáticas". No estoy tan de acuerdo en la utilidad del mismo como texto básico de álgebra en una licenciatura de matemáticas. (E.E.)

ANCIENT EGYPTIAN SCIENCE, A SOURCE BOOK, por Marshall Clagett. American Philosophical Society; Philadelphia, 1989.

Por si les faltaba razón y ejemplos a quienes, entre nosotros, consideran eutanásica, si no de depuración política, la norma sobre edad de jubilación de los profesores, he aquí que Marshall Clagett, septuagenario, nos ofrece una obra maestra de egiptología, la primera, además, que se ocupa in extenso de la ciencia en el país del Nilo.

Clagett entró hace años en el panteón de fama con sus estudios sobre la física medieval y, sobre todo, por sus trabajos sobre la prolongación de la geometría arquimediana en Occidente. Si el lector echa un vistazo al libro homenaje que le dedicaron amigos y discípulos con motivo de su jubilación académica (*Mathematics and its applications in science and natural philosophy in the Middle Ages*, Cambridge, 1987), observará que aparecen, sólo de refilón, y a modo de promesa de futuro, sus aficiones egiptológicas.

No es, sin embargo, esta obra un escrito de urgencia, de quien sospecha cercano el hito de arribada. El Clagett minucioso de siempre, puntilloso en la interpretación correcta del texto, resurge con el brío de años atrás. Nadie diría que es un advenedizo. La verdad es que sólo lo es en punto a redacción. Se adivina que hay muchas notas, muchos viajes, muchas fotografías, largas noches de contrastación de textos lupa en ristre, muchas lecturas también, detrás de esa exposición paladina de lo que dicen y no pueden decir unos jeroglíficos de endiablada polisemia. Bagaje necesario para unir la interpretación de esta tumba con la de aquella estela y con la tablilla de más allá, para enhebrar una historia verosímil, quizá veraz. Los dos primeros tomos que componen este primer volumen abarcan dos capítulos: los frutos de la actividad de los escribas y el mundo y su creación. A esos responde también el subtítulo "Conocimiento y orden", substanciados, a su vez, en dos términos —*rekh* y *maat*—. El vocablo *rekh* compendia el ideal del hombre ilustrado, conocedor del secreto de la escritura y, con él, del arte de medir, contar e historiar. La voz *maat* explicita el orden que reina en el mundo, su origen y gobierno. Clagett sabrá eludir, a lo largo de su exposición, la hipóstasis de esos términos, al estilo de lo que ha venido ocurriendo con cierta historia de la ciencia griega que se recrea en la metafísica, o al menos en la metaciencia, de los *lógoi, kósmos, téchne,* etcétera.

Clagett prefiere revelarnos el concepto encerrado en *rekh* detallando el desarrollo de la escritura jeroglífica, sus formas, sus contenidos; es decir, nos recrea poco menos que la vertiente social, política y económica de tres mil años de historia. Para ello, aporta, como fuente, nueve documentos en este primer capítulo, traducidos por él, o al menos retocados si existe una versión más o menos fiel. Las notas que acompañan suelen ser un análisis crítico de otras opiniones o una exposición detallada, y nueva, del significado del texto original. Vale la pena ponderar su esfuerzo de síntesis en un segmento de la cultura en el que apenas la cronología dinástica es lo único más o menos asentado.

A riesgo de simplificar, el tránsito del mito a la razón se da en el Egipto faraónico en el paso de politeísmo, o de dioses locales, al monoteísmo centralizado. Lo cierto, sin embargo, es que no podemos separar la filosofía natural o la física del mito y la magia, porque no existe, además, un pensamiento precientífico único, ni una sola cosmología. En efecto, el sistema heliopolitano de la creación del mundo difiere, por ejemplo, de la cosmogonía de Hermópolis, y ésta de la menfítica, que a su vez lo hace de las cosmogonías tebanas de Amón-Ra, de Aten, del templo de Esna... Cosmogonías dispares, e incluso con variantes en su seno, que compartían algunos rasgos unificadores: rechazo de la creación *ex nihilo* con la preexistencia de un abismo originario, extrapolación cósmica de la inundación periódica del Nilo (las aguas como fuente de vida) y exposición de la denominación creadora, es decir, las cosas comienzan a existir cuando se les pone un nombre.

Tras la cosmogonía, la cosmología, que Marshall aborda considerando primero las propiedades del cosmos, los seres que lo inhabitan —visibles e invisibles— y la naturaleza de las fuerzas que lo mantienen en la existencia y dan cuenta de su desarrollo armónico.

La segunda parte de la cosmología, la astronomía, constituye el objeto del volumen dos. Extrae, del análisis de distintos "libros" del hontanar egipcio, las constantes posibles de un cielo dividido en doce horas, las que dura el viaje de los muertos hacia el más allá. Cierra esta primera entrega de lo que promete ser la obra cumbre de la ciencia egipcia —como la escrita por Joseph Needham sobre China— con varios apéndices: cronología, bibliografía e índices, amén de un centenar largo de páginas iconográficas, la mayoría concisamente comentadas. (L.A.)

Apuntes

Hay muchas formas de abordar la naturaleza viva. La simulación, en el ordenador, de su comportamiento es hoy la más socorrida. Los químicos de la vieja escuela prefieren una vía más tangible: remedarla en el tubo de ensayo, que produce especial satisfacción cuando se tiene éxito. Por ejemplo, si se logra imitar el ensamblaje espontáneo de una cadena en doble hélice partiendo de precursores lineales. Hazaña que ha conseguido Jean-Marie Lehn: ha sintetizado una clase de sistemas artificiales, los nucleohelicatos, que poseen una estructura en doble hélice y presentan la interacción selectiva de las bases nucleotídicas.

Las galaxias espirales, y no sólo la nuestra, puede que no sean tan transparentes como habitualmente se piensa. Se están recogiendo datos que apuntan más bien a que dentro de ellas abunda el polvo interestelar, lo que las haría bastante opacas a nuestra vista, salvo en lo referente a su parte más externa. De confirmarse, habría que repensar algunas cuestiones cosmológicas, como la de la famosa materia oculta del universo, las dimensiones de la parte observable de éste o la gran luminosidad de las galaxias espirales en el infrarrojo.

A pesar de ser el resultado de varios lustros de cuidadoso trabajo por parte de centenares de los mejores investigadores y técnicos del mundo, el telescopio espacial Hubble es defectuoso. Mientras las primeras dificultades experimentadas tras su lanzamiento podrían disculparse por la novedad y complejidad del propósito, se ha llegado a la conclusión de que hay un fallo mucho más elemental y perturbador: el sistema de espejos del telescopio, aun tallado y pulido con la mayor precisión del mundo, presenta aberración de esfericidad, una de las más viejas pesadillas de los fabricantes de instrumentos ópticos. Alguno de los espejos, probablemente el primario, no tiene la curvatura debida. Esto impide enfocarlo con la precisión necesaria y hará que sus imágenes no sean mucho mejores que las obtenidas con los más elaborados instrumentos terrestres. Al menos hasta que no se encuentre un remedio, que podría ser el clásico de colocarle una lente correctora, lo que debería hacer un astronauta y no sucedería antes de 1993.

Por cristalografía de rayos X se ha determinado la estructura tridimensional de la lipasa pancreática, enzima de importancia principal en la absorción de grasas de la dieta; hidroliza los triglicéridos en diglicéridos y, luego, en monoglicéridos y ácidos grasos libres. Se trata de una glicoproteína monocatenaria de 449 aminoácidos, cuya estructura primaria se ha establecido por secuenciación de los clones de ADN complementario. Gracias a ese trabajo se sabe que el aminoácido serina de la posición 152 resulta esencial para la acción catalítica de la enzima.

Los núcleos celulares fraccionan las moléculas que vienen del citosol: macromoléculas como proteínas y ARN, pequeños péptidos, aminoácidos, azúcares y ciertos iones. El núcleo las acumula o las torna a expulsar. Ese comportamiento se ha explicado mediante un doble mecanismo: retención selectiva de las moléculas en libre difusión y franqueo, asimismo selectivo, de la entrada a través de la membrana nuclear. Admitíase que la membrana se limitaba a cerrar el paso de las macromoléculas; ahora se propone la existencia de canales selectivos de iones potasio, que presentarían múltiples estados de conductancia o acarreo.

Los diamantes no constituyen ninguna rareza. Se han descubierto en restos carbonáceos de explosiones, en meteoritos, nucleados en fase gaseosa y desarrollados en forma laminar en hidrógeno. Aunque las condiciones varían en los cuatro casos, el tamaño de los diamantes viene a ser casi igual: de escasos nanómetros de diámetro. La posibilidad de que abunden más, aunque por sus dimensiones pasen inadvertidos, se apoya en sus energías de superficie, muy importantes a la hora de estabilizar los diamantes microcristalinos. A tenor de los enlaces superficiales con los átomos de hidrógeno, los diamantes cuyo diámetro no llega a los tres nanómetros se ven energéticamente privilegiados sobre los precursores del grafito.

Para saciar el hambre de los pueblos no basta a veces enseñarles a pescar, es preciso urgirles a buscar su propia caña o tejer sus redes. La arqueoagricultura acaba de poner de manifiesto que entre el 50 y el 75 por ciento de las feraces terrazas incas se hallan hoy abandonadas y se dedicaba al regadío un 40 por ciento más de los valles costeros del Perú antes de la conquista. Los habitantes precolombinos de los Andes poseían técnicas agrarias idóneas para sacarle el máximo rendimiento al relieve: las gochas o campos en círculo de los altos andinos, jardines en los bajíos y campos acompasados a un refinado sistema de canalización que discurría por la cordillera.

Seguiremos explorando los campos del conocimiento



INFECCIONES ASOCIADAS AL SIDA, por John Mills y Henry Masur
Mientras no se descubra una vacuna contra el sida, la esperanza de prolongar la vida del paciente descansará en refinar la terapia de las infecciones oportunistas del virus.

EL LENGUAJE DE LOS FRACTALES, por Hartmut Jürgens, Heinz-Otto Peitgen y Dietmar Saupe
Estas estructuras detalladísimas hasta lo increíble son algo más que meras curiosidades matemáticas. La geometría fractal describe sucintamente objetos y procesos naturales complejos.

ANTICUERPOS MONOCLONALES DE SEGUNDA GENERACION, por Jorge V. Gavilondo Cowley
Elementos esenciales para el diagnóstico y la terapia de enfermedades agudas y crónicas, los anticuerpos monoclonales se han convertido en poderosa herramienta de la investigación básica.

TENDENCIAS HACIA EL CALENTAMIENTO GLOBAL, por Philip D. Jones y Tom M.I. Wigley
El análisis de los archivos terrestres y marítimos confirma que nuestro planeta se ha calentado medio grado Celsius en el último siglo. No obstante, las tendencias hacia el calentamiento en el futuro siguen siendo inciertas.

EL ARTE MAYA A REGISTRO, por June Kinoshita
La aceleración del desarrollo económico está echando a perder los murales mayas, por lo que las réplicas de éstos pintadas a mano vienen a ser fuentes imprescindibles para la arqueología.

CUANDO LOS PUNTOS DE FUSION Y CONGELACION NO SON LOS MISMOS, por R. Stephen Berry
Los agregados atómicos ofrecen pistas para entender los fenómenos de solidificación y fusión. Estos agregados, que contienen desde unos pocos a varias decenas de átomos o moléculas, pueden coexistir en las fases sólida y líquida.

SEMILLAS DISPERSADAS POR HORMIGAS, por Steven N. Handel y Andrew J. Beattie
Miles de especies vegetales dependen de las hormigas para dispersar sus semillas. Con señuelos alimenticios especiales y otras adaptaciones, una planta puede inducir a los insectos para que transporten las semillas sin dañarlas.

LOS SUPERCONDUCTORES MAS ALLA DEL 1-2-3, por Robert J. Cava
El óxido de itrio1-bario2-cobre3 es sólo uno de los muchos superconductores nuevos de alta temperatura. Todos ellos contienen planos formados por átomos de cobre y oxígeno que constituyen una "autopista" perfecta para los electrones.

